

„Qualitätsbeurteilung der Messung systolischer Zeitintervalle  
in der neonatologischen Praxis durch Analyse der intra- und  
interindividuellen Reliabilität“

*Prospektive klinische Studie*

Dissertation  
zur Erlangung des akademischen Grades

Doctor Medicinae (Dr. med.)

An der Medizinischen Fakultät  
der Universität Leipzig

eingereicht von  
Riekje Isberner  
Geboren 14.09.1994 Berlin

angefertigt in Klinik und Poliklinik für Kinder- und Jugendmedizin  
Abteilung für Neonatologie  
der Universität Leipzig

Betreuer:  
Prof. Dr. Matthias Knüpfer

Beschluss über die Verleihung des Doktorgrades vom: 21.09.2021

## Widmung

Für Malte und Nane, meine Geschwister.

Euretwegen bin ich Kinderärztin geworden.

# Inhaltsverzeichnis

1. Abkürzungsverzeichnis .....	III
2. Einführung.....	1
2.1 Herzmechanik .....	1
2.2 Herz- und Kreislauffunktion in der Neonatologie.....	2
2.3 Systolische Zeitintervalle .....	4
2.3.1 Allgemeines.....	4
2.3.2 Bedeutung systolischer Zeitintervalle in der Diagnostik des hsPDA.....	7
3. Aufgabenstellung .....	9
4. Material/Methodik.....	10
4.1 Patientenkollektiv .....	10
4.2 Datenerhebung und Dokumentation .....	11
4.3 Theorie und Durchführung der Echokardiographie zur STI-Ermittlung .....	14
4.4 Grundlagen zur Statistik .....	16
4.4.1 Allgemeine Statistik.....	16
4.4.2 Bland et Altmann Plots.....	16
4.4.3 Repeatability Coefficient and Repeatability Index .....	18
4.4.4 Cohen's Kappa coefficient .....	20
5. Ergebnisse .....	24
5.1 Basisdaten .....	24
5.2 Übersicht der Ergebnisse der Messungen der STIs .....	26
5.2.1 Untersuchungsvoraussetzungen .....	26
5.2.2 Übersicht zur durchschnittlichen Dauer der erhobenen STIs und zu den ermittelten WR28 .....	28
5.3 Betrachtung der Daten mittels Bland et Altmann plots.....	29
5.4 Inter-/Intraindividuelle Wiederholbarkeit anhand des RC und dem RI.....	32
5.5 Inter-/Intraindividuelle Wiederholbarkeit anhand des RC und dem RI bei gleichbleibenden Voraussetzungen hinsichtlich HF und Agitiertheit.....	38
5.6 Inter-/Intraindividuelle Wiederholbarkeit anhand des Cohens Kappa Koeffizienten.....	46
5.7 Inter-/Intraindividuelle Wiederholbarkeit anhand des Cohens Kappa Koeffizienten bei gleichbleibenden Voraussetzungen hinsichtlich HF und Agitiertheit .....	54
5.8 Zusammenfassung der wichtigsten Ergebnisse .....	61
6. Diskussion .....	62
7. Zusammenfassung .....	71
8. Abbildungsverzeichnis .....	76

<b>9. Tabellenverzeichnis.....</b>	<b>77</b>
<b>10. Literaturverzeichnis .....</b>	<b>79</b>
<b>11. Anhang .....</b>	<b>83</b>
<b>12. Selbstständigkeitserklärung .....</b>	<b>85</b>
<b>13. Lebenslauf.....</b>	<b>86</b>
<b>Danksagung.....</b>	<b>89</b>

## 1. Abkürzungsverzeichnis

APGAR	Apgar-score (engl. Appearance, Pulse, Grimace, Activity, Respiration)
BNP/NT-proBNP	B-type natriuretisches Peptid/ N-terminales proBNP
EF	Ejektionsfraktion (engl. Ejection fraction)
EKG	Elektrokardiogramm
EL	Erfahrungslevel
ELBW	Neugeborenes mit Geburtsgewicht < 1000g (engl. Extremely low birth weight)
FG	Frühgeborenes
GA	Gestationsalter (engl. Gestational age)
GG	Geburtsgewicht
HF	Herzfrequenz
hsPDA	Hämodynamisch signifikanter persistierender Ductus arteriosus (engl. hemodynamically significant PDA)
HMV	Herzminutenvolumen
ICG	Impedanzkardiographie (engl. Impedance cardiography)
KA	Katecholamine
KI	Konfidenzintervall
LA/Ao ratio	Verhältnis linkes Atrium zur Aortenwurzel (engl. Left atrial/aortic root ratio)
LPEP	Linksventrikuläre Präejektionsphase (engl. Left ventricular pre-ejection period)
LT	Lebenstag
Lv	Linksventrikulär
LVEF	Linksventrikuläre Ejektionsfraktion
LVET	Linksventrikuläre Ejektionszeit (engl. Left ventricular ejection time)
LWR	Linksventrikuläre Weissler ratio
NG	Neugeborene
PCG	Phonokardiogramm (engl. Phonocardiogram)
PDA	Persistierender Ductus arteriosus Botalli
PEP	Präejektionsphase (engl. Pre-ejection period)
PPH	Persistierende pulmonale Hypertonie
PPG	Photoplethysmographie (engl. Photoplethysmography)

PW-Doppler	PW-Doppler (engl. Pulsed wave – Doppler)
RC	Wiederholbarkeitskoeffizient (engl. Repeatability coefficient)
RCT	Randomisiert kontrollierte Studien (engl. Randomised controlled trials)
RI	Wiederholbarkeitsindex (engl. Repeatability index)
RPEP	Rechtsventrikuläre Präejektionsphase (engl. Right ventricular pre-ejection period)
Rv	Rechtsventrikulär
RVET	Rechtsventrikuläre Ejektionszeit (engl. Right ventricular ejection time)
RWR	Rechtsventrikuläre Weissler ratio
SD	Standardabweichung (engl. Standard deviation)
SST	Schwangerschaftstage
SSW	Schwangerschaftswoche
STI	Systolische Zeitintervalle (engl. Systolic time intervals)
VAS	Visuelle Analogskala
VET	Ventrikuläre Ejektionszeit (engl. Ventricular ejection time)
VLBW	Neugeborenes mit Geburtsgewicht <1500g (engl. Very low birth weight)

## 2. Einführung

### 2.1 Herzmechanik

Das Herz gewährleistet über seine Pumpfunktion die Versorgung des Körperkreislaufs mit Sauerstoff angereichertem Blut aus dem Lungenkreislauf. Die Funktion aller lebenswichtigen Organe hängt somit u.a. von der Erbringung der nötigen Pumpleistung des Herzens ab. Die Kontraktionsphase, sowie die Entspannungsphase des Myokards eines jeden Herzens sind reguliert durch die Regeneration von Calciumionen.<sup>1</sup> Es besteht ein ausgeprägter Zusammenhang zwischen der Zeitabfolge beider Abschnitte des Herzzyklus und der Vitalität der Herzmuskelzellen. Diese muss gewährleistet sein, um die Blutversorgung der lebenswichtigen Organe, entsprechend ihres Metabolismus, sicherzustellen.<sup>2</sup> Über einen elektrischen Reiz, generiert von Schrittmacherzellen, entsteht ein Aktionspotenzial. Der primäre Schrittmacher ist der Sinusknoten. Über das Reizleitungssystem des Herzens wird die geordnete Übertragung der Erregung auf alle Herzmuskelzellen sichergestellt. Es folgt die elektromechanische Kopplung - die Transformation des elektrischen Signals in die mechanische Bewegung von Aktin- und Myosinfilamenten. Das Herz kontrahiert.

Der Herzzyklus setzt sich aus Systole (Kontraktionsphase) und Diastole (Entspannungsphase) zusammen, die jeweils erneut in 2 Perioden unterteilbar sind. Die Systole gliedert sich in eine Anspannungs- und eine Austreibungsphase, die Diastole umfasst die Entspannungs-, sowie die Füllungsphase. Die rhythmische Abfolge der einzelnen Abschnitte sichert die adäquate Pumpwirkung. Dabei nimmt die Systole  $\frac{1}{3}$ , die Diastole  $\frac{2}{3}$  der Zeit eines Zyklus in Anspruch.

## 2.2 Herz- und Kreislauffunktion in der Neonatologie

Kardiovaskuläre Pathologien und Kreislaufdysregulationen sind bei Neugeborenen (NG) häufig zu beobachten und stellen für Neonatologen eine große diagnostische und therapeutische Herausforderung dar.

Die regelrechte postnatale Kreislaufumstellung ist entscheidend für das Überleben der NG. Angeborene Herzfehler, die persistierende pulmonale Hypertonie (PPH) und persistierende kardiovaskuläre/kardiale Shunts oder auch eine Neugeborenenensepsis stellen nur beispielhaft zu erwähnende Herausforderungen dar, hinsichtlich einer adäquaten hämodynamischen Regulation.<sup>3</sup>

NG und Frühgeborene (FG) bilden ein besonderes Patientengut ab, das Herz-Kreislaufsystem unterscheidet sich grundlegend von dem, älterer Kinder und Erwachsener. Das Herz eines NG verstoffwechselt im Vergleich zu einem adulten Herz beispielsweise vermehrt Laktat und Glukose.<sup>4</sup> Das Myokard der NG enthält außerdem deutlich mehr Mitochondrien und zu einem größeren Anteil nicht-kontraktile Elemente. Weiterhin ist die Calcium-induzierte-Calciumfreisetzung des sarkoplasmatischen Retikulums aufgrund der Unreife der Myofibrillen insuffizient.<sup>3</sup> Folglich ist die kontraktile Reserve der NG begrenzt, sodass beispielsweise ein Anstieg der Nachlast schlechter toleriert wird und zu einer verminderten Auswurfleistung führt.<sup>5</sup>

Häufig zu beobachtende Symptome einer Dysregulation sind eine kompensatorische Tachykardie und im Verlauf ein erniedrigter Blutdruck. Ein Laktatanstieg mit Ausbildung einer metabolischen Azidose im Rahmen der insuffizienten Versorgung der Organe mit Sauerstoff und auch Oligurie oder Anurie im Rahmen eines Schockgeschehens sind dann möglicherweise zu beobachten. Das hämodynamische Monitoring war lange Zeit auf eine kontinuierliche Messung der Herzfrequenz (HF), des Blutdrucks und der Sauerstoffsättigung in den Arteriolen beschränkt, sowie auf die intermittierende Beurteilung klinisch-chemischer Parameter der Organperfusion, wie Diurese, Rekapillarisierungszeit und die Bestimmung des Serumlaktats.<sup>6,7</sup>



Rein klinisch lassen sich jedoch die Ätiologien kardiovaskulärer Dysregulationen kaum unterscheiden. Ein zu niedriger Blutdruck bei NG kann exemplarisch sowohl in einem Volumenmangel, als auch in einer schlechten Ventrikelkontraktion, einem distributiven Schock (z.B. septischer Schock), einem persistierenden Ductus arteriosus (PDA) oder anderen Ursachen begründet liegen.

Genormte Werte bezüglich des Blutdrucks zur Beurteilung der Herz-Kreislaufsituation lassen sich, insbesondere bei FG, aufgrund unterschiedlicher Gegebenheiten hinsichtlich der Reife des kardiopulmonalen Systems und auch der Unreife anderer Organsysteme schwer erheben.<sup>6</sup> Die Definition einer Hypotonie beim NG ist folglich uneinheitlich. In der Literatur finden sich beispielsweise Einteilungen anhand der Blutdruckwerte unterhalb der fünften bzw. zehnten Perzentile des jeweiligen Gestationsalters/Alters. Nach dieser Definition würden 50% der Very low birth weight-Kinder (VLBW) auf den neonatologischen Intensivstationen an postnataler Hypotonie leiden.<sup>6,7</sup> Weiterhin besteht die Einteilung eines zu niedrigen Blutdrucks in drei Level anhand des zerebralen Blutflusses (korreliert zum MAP).<sup>8</sup> Eine Hypotonie wird hierbei definiert als ein mittlerer Blutdruck von 28-30 mmHg für ELBW und 22-24 mmHg für VLBW.<sup>6,9</sup>

Beispielhafte Ansätze zur Optimierung der pathophysiologischen Ergründung hämodynamischer Dysregulationen stellen die funktionelle Echokardiographie, das Nicht-invasive-cardiac-output-monitoring, die Impedanzkardiographie (ICG) und auch die Nahinfrarotspektroskopie dar.<sup>6,7,10-13</sup> Hierbei gilt die funktionelle Echokardiographie weiterhin als meistgenutztes Tool, um nicht nur die globale Herzfunktion, einschließlich der kardialen Füllung, der Kontraktilität und des Auswurfs, sondern auch die kardialen Strukturen zu beurteilen. Sie ist somit eine hilfreiche Methode, um die Pathophysiologie der kardialen Insuffizienz zu definieren. Die echokardiographische Ermittlung der Ejektionsfraktion (EF) bzw. des Herzminutenvolumens (HMV) wird diagnostisch in der Mehrheit der klinischen Zentren angewandt. Eine weitere Möglichkeit, nicht nur zur Abschätzung des HMV, aber auch zur Beurteilung der ventrikulären Kontraktilität stellt die dopplersonographische Erhebung systolischer Zeitintervalle dar (siehe 2.3).

## 2.3 Systolische Zeitintervalle

### 2.3.1 Allgemeines

Systolische Zeitintervalle (STI), die Präejektionsphase (PEP) und die ventrikuläre Ejektionszeit (VET) einschließend, dienen der Beurteilung der rechts- bzw. linksventrikulären (rv/lv) Funktionalität. Die echokardiographische Messung der STIs stellt eine nicht-invasive Möglichkeit dar, die globale Herzfunktion zu ermitteln.<sup>1,2,14,15</sup>

Die der Ejektionsphase vorausgehenden PEP umfasst die elektromechanische Verzögerung und die isovolumetrische Anspannungsphase.<sup>14</sup> Sie beginnt mit dem Start der Ventrikeldepolarisation (Q-Zacke im Elektrokardiogramm [EKG]) und endet mit dem Moment der Aorten- bzw. Pulmonalklappenöffnung, die die Austreibungsphase einleitet (siehe Abbildung 1). Die lv Ejektionsphase oder „Left ventricular ejection time“ (LVET) ist begrenzt durch Öffnung und Schluss der Aortenklappe, die „Right ventricular ejection time“ (RVET) entsprechend durch das Öffnen und Schließen der Pulmonalklappe.<sup>2,15,16</sup>

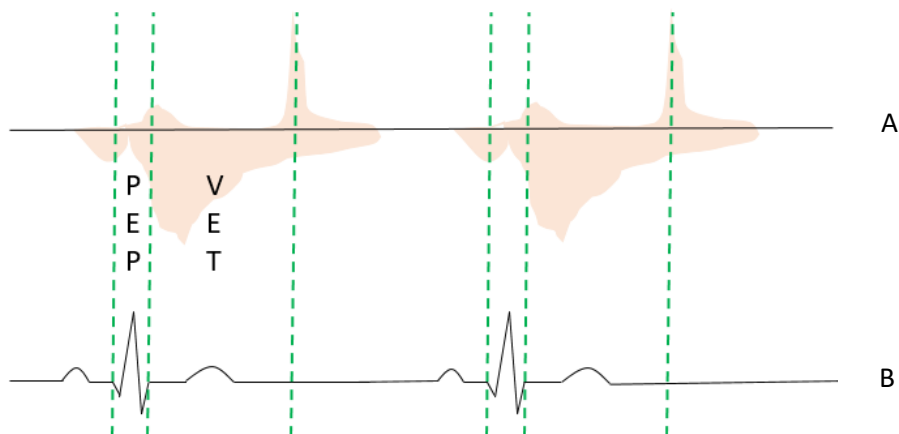


Abbildung 1 Skizze PEP bzw. VET-Messung mit Hilfe von Flussprofil (A) und EKG (B); siehe vergleichend Abbildung 2 unter 4.3

Das Schlagvolumen und somit auch das HMV stehen in engem Zusammenhang mit der LVET. Ist das Zeitintervall verlängert, vergrößert sich in der Regel auch das HMV. Ein gesundes Herz zeigt eine kurze PEP und eine lange LVET, während eine verminderte Funktionalität/Dysfunktion des Myokards die PEP verlängert und die VET verkürzt.<sup>1,2,14,16–20</sup> Die einfache echokardiographische Ermittlung der PEP ermöglicht Rückschlüsse auf die Herzmuskelkontraktion. Die PEP ist jedoch nicht nur abhängig von der Kontraktilität des Myokards, sondern auch von der Vor- und Nachlast.<sup>16</sup> Bei Patienten, die an arterieller Hypertonie leiden, wurde beispielsweise eine verlängerte PEP ermittelt.<sup>20,21</sup> Auch die adäquate elektrische Aktivität und die damit verbundene Reizleitung des Herzens hat Einfluss auf die Dauer der PEP.<sup>16,22</sup>

Es gibt zahlreiche Pathologien, die mit einer myokardialen Dysfunktion einhergehen und eine Verlängerung der PEP verursachen, beispielsweise Herzrhythmusstörungen (Linksschenkelblock etc.), aber auch eine verminderte Vorlast, so wie negativ inotrop wirkende Mittel. Eine Verlängerung der LVET liegt z.B. häufig in Erkrankungen der Aortenklappe begründet und kann hinweisend auf eine relevante Aortenklappenstenose sein.<sup>23</sup> Diese Pathologien spielen bei Neugeborenen eine untergeordnete Rolle. Eine Verlängerung der LVET ist in der Neonatologie primär ein Hinweis auf eine Volumenüberlastung, am ehesten im Rahmen eines hämodynamisch signifikanten PDA (hsPDA).<sup>19</sup>

Ebenso gibt es Faktoren, welche die STIs verkürzen. Eine verkürzte PEP kann man im Rahmen einer Therapie mit positiv inotrop wirkenden Medikamenten, sowie einem erniedrigten lv Druck, aber auch bei Erkrankungen der Aortenklappe beobachten. Die LVET verkürzt sich ebenfalls unter der Gabe von positiv, allerdings auch negativ inotrop wirkenden Agenzien. Weitere relevante Einflüsse hinsichtlich einer Verkürzung der LVET sind eine bestehende lv Dekompensation oder eine verringerte Vorlast. Eine bestehende Hypovolämie kann sowohl die LVET, als auch die RVET verkürzen.<sup>15,23</sup>

STIs korrelieren invers mit der HF. Mit steigender HF nimmt vor allem die Dauer der VET ab, die PEP zeigt nur eine geringgradige Anpassung.<sup>2,23,24</sup>

Die Dauer der STIs steigt von der Kindheit bis zur Pubertät an, insbesondere die PEP.<sup>23</sup> Die VETs bei Frauen sind generell länger als bei Männern.<sup>2</sup>

Die nicht-invasive Erhebung der STIs erfolgt in unserem Zentrum durch die echokardiographische Untersuchung der Neugeborenen mit der Doppler Methode (siehe 4.3) bei gleichzeitiger Ableitung eines EKG. Anhand der Flusskurve über der Aortenklappe (darstellbar in kurzer parasternaler Achse) und dem EKG können die lv STIs ermittelt werden. Analog verhält es sich mit den rv STIs über der Pulmonalklappe.<sup>15,19</sup> Bei Erwachsenen werden STIs häufig auch durch die parallele Ableitung eines EKGs oder eines Phonokardiogramms (PCG) und der invasiven Carotiden-Puls-Messung ermittelt.<sup>14,22,23</sup>

Ein Index, um die kardiale Funktion zu normieren, ist die sogenannte Weissler ratio (WR), oder der Weissler Quotient.

<b>Weissler ratio = <math>\text{PEP}/\text{VET}</math> <sup>1</sup></b>
---

Der Wert sollte bei NG in einem Normbereich von 0,35-0,42 liegen.<sup>15</sup> Er ist weniger von der HF abhängig und somit ein besserer Parameter für die Beurteilung der gesamten kardialen Funktion bzw. des hämodynamischen Status als die PEP oder die VET allein.<sup>1,16,22</sup>

Ein erhöhter Quotient der rv STIs gilt am ersten Lebenstag (LT) als physiologisch, da noch ein erhöhter pulmonaler Widerstand besteht und folglich die rv Präejektionsphase (RPEP) verlängert ist. Sollte sich der Wert im Verlauf nicht verringern oder sogar ansteigen, kann dies ein Hinweis auf eine PPH sein, den es weiter abzuklären gilt, insbesondere, wenn zeitgleich ein erhöhter Sauerstoffbedarf besteht.<sup>15,25</sup>

Ist die lv WR (LWR) erhöht, ist dies ein möglicher Hinweis auf eine verminderte Funktion des Myokards, die es ebenso zu spezifizieren gilt. Ist sie hingegen erniedrigt, ist dies hinweisend auf einen PDA. Pathophysiologisch erklärbar ist dies durch die ventrikuläre Volumenüberladung und eine konsekutive Verlängerung der LVET, sowie Verkürzung der PEP, aufgrund des Frank-Starling Mechanismus.<sup>19</sup>

### 2.3.2 Bedeutung systolischer Zeitintervalle in der Diagnostik des hsPDA

Der Goldstandard hinsichtlich der Diagnose eines PDA ist nach wie vor die Echokardiographie.<sup>26,27</sup> Bei bis zu einem Drittel aller FG mit einem Geburtsgewicht (GG) <1500 g (VLBW) ist ein PDA diagnostizierbar.<sup>28</sup> In der Kohorte der Extremely low birth weight-Kinder (ELBW), NG mit einem GG unter 1000 g, kann man davon ausgehen, dass 55% bis nahezu zwei Drittel betroffen sind (65%).<sup>29,30,31</sup>

Wie im vorherigen Kapitel beschrieben, weist eine erniedrigte LWR auf einen PDA hin. Dabei kann man, sofern eine signifikant verkürzte PEP bei gleichzeitig signifikant verlängerter LVET vorliegt, von einer Spezifität von 100% ausgehen, hinsichtlich der Diagnose einer hämodynamischen Relevanz des PDA.<sup>15</sup> Der Bestimmung der STI werden dahingehend signifikante Vorteile gegenüber anderen sonographischen Methoden, beispielsweise der Bildung der Ratio aus linkem Vorhof- und Aortendurchmesser (s.u.), in der Beurteilung eines hsPDA zugeschrieben.<sup>15,19</sup>

Eine weitere häufig genutzte sonographische Methode zur Eruierung der hämodynamischen Signifikanz des PDA ist die Bestimmung der left atrial/aortic root ratio (LA/Ao-ratio), eine Gegenüberstellung des Durchmessers des linken Vorhofes gegenüber dem Aortenwurzeldurchmesser. Andere dopplersonographische Methoden zur Beurteilung der Relevanz des PDA, beispielsweise die Ermittlung des Resistance Index in der Arteria cerebri anterior und dem Truncus coeliacus, sollen ebenfalls eine hohe Wiederholbarkeit aufweisen.<sup>31</sup> Weiterhin ist die Messung der Flussgeschwindigkeit im aufsteigenden Ast der Aorta und in der Pulmonalarterie möglich, woraufhin die Bildung von Geschwindigkeits-Zeit-Integralen zur Einschätzung des PDA erfolgt.

Klinische Hinweise für einen hsPDA sind kardiale (Kardiomegalie und aktives Präkordium; Systolikum/Diastolikum; „Maschinengeräusch“ etc.), aber auch respiratorische Symptome (zunehmende Abweichungen in der Blutgasanalyse, Bedarf Atemhilfe/Sauerstoffgabe). Weiterhin sind den Kreislauf betreffende klinische Zeichen (verlängerte Rekapillarierungszeit, starke periphere Pulse etc.) und renale Symptome (Oligurie) registrierbar, sowie das Ansteigen bestimmter Biomarker (B-type natriuretisches Peptid [BNP], N-terminales proBNP [NT-proBNP]).<sup>19,27,32</sup>

Ein hsPDA spielt eine große Rolle, da er konsekutiv nicht nur zu einer übersteigerten pulmonalen Perfusion, sondern auch zu einer renalen, intestinalen und zerebralen Minderperfusion führen kann.<sup>19,33,34</sup> Die nekrotisierende Enterokolitis, sowie die periventrikuläre Leukomalazie und die bronchopulmonale Dysplasie stellen gefürchtete Komplikationen dar.<sup>14,19,32,33</sup> Ebenso wird das gehäufte Auftreten metabolischer Azidosen, eine längere Beatmungsabhängigkeit und eine schlechtere Gewöhnung an das Stillen beschrieben.<sup>32</sup> Ein PDA geht bei FG mit einer erhöhten Morbidität und Mortalität einher.<sup>31,32</sup> Dabei ist das Risiko eines hsPDA bei ELBW, mit einer Wahrscheinlichkeit von 50-70% am höchsten.<sup>26,35-37</sup>

Zonnenberg et al veranschaulichten 2012 in ihrem Review aus 67 randomisierten Studien (RCTs), dass ein hsPDA an verschiedenen Zentren unterschiedlich definiert wird.<sup>38</sup>

Um die Versorgungsqualität der Patienten auch zentrumsübergreifend gewährleisten zu können, wäre es wünschenswert, eine standardisierte Untersuchungsmethode mit klar definierten Kriterien für die häufigsten Pathologien festzulegen. Beispielsweise einen internationalen Konsensus für einen hsPDA, auch um die Vergleichbarkeit von Studien zu gewährleisten.<sup>32</sup> Wichtig wäre demnach eine wiederholbare, also verlässliche Messmethode zur Definition eines diagnostischen Goldstandards. Die Ermittlung der STI könnte dabei eine Rolle spielen.

### 3. Aufgabenstellung

Evidenzbasierte Medizin erfordert standardisierte Methoden zur Ermittlung klinischer Daten, deren Interpretation entscheidend sein kann bei der Einleitung diverser Diagnostiken und Therapien. Die Wiederholbarkeit der Untersuchungsmethode muss garantiert sein, um die Aussagekraft der erhobenen Werte unabhängig vom Untersuchungszeitpunkt/Untersucher zu gewährleisten und eventuelle Einflussfaktoren zu eruieren.

Im vorherigen Kapitel wurde die Bedeutung der echokardiographischen Untersuchung der NG dargestellt. Zweifellos ist die Echokardiographie eine wichtige Untersuchung zur Einschätzung der hämodynamischen Situation im Bereich der Neonatologie. Über die beste Methode zur Diagnostik gibt es bisher keinen Konsens. Die Messung der STIs ist einfach und die Daten sind nicht-invasiv schnell zu erheben. Wenn diese einfache Methode propagiert und angewendet werden soll, ist es entscheidend, zu klären, wie verlässlich, also reproduzierbar, die Daten erhoben werden können. Da dem Perinatalzentrum Level 1 der Universitätsklinik Leipzig die routinemäßige Durchführung eines Herzultraschalls bei jedem NG erfolgt und im Rahmen dessen STIs zur Bewertung des kardialen Status der Patienten verwendet werden, wurde in der vorliegenden Arbeit die Wiederholbarkeit (repeatability) und die Verlässlichkeit (reliability) der Messung dieser Parameter untersucht.

Dabei wurde auch analysiert, inwiefern die Erfahrung des sonographierenden Arztes einen Einfluss auf die Reproduzierbarkeit der Ergebnisse hat. Weiterhin wurden zahlreiche zuvor definierte Patientencharakteristika hinsichtlich ihrer Bedeutung für die Messergebnisse untersucht. Es wurden intra- und interindividuelle Unterschiede der erhobenen Messwerte analysiert. Anhand des Ausmaßes der Abweichungen und der ggf. klinisch differenzierten Bewertung des Status eines Kindes wurde die Reproduzierbarkeit der ermittelten Daten und damit die zukünftige Bedeutung der STIs im klinischen Alltag eruiert.

Es wurde im Verlauf der Untersuchungen sehr streng auf die Vergleichbarkeit der Untersuchungsbedingungen geachtet. Dabei wurden Veränderungen der Vigilanz (ermittelt anhand visueller Analogskala = VAS) oder der HF zwischen den einzelnen Untersuchungen besonders registriert.

Wir erhoffen uns Aussagen zu der wichtigen Frage, wie anfällig die Messwerte im intraindividuellen und interindividuellen Vergleich sind, um die Stabilität der Messwerte und folglich ihre Verwertbarkeit für die Diagnostik einschätzen zu können.

## 4. Material/Methodik

### 4.1 Patientenkollektiv

Im Zeitraum vom 06.10.2016 bis zum 09.02.2018 wurden alle Kinder, die in der klinischen Arbeit routinemäßig oder indikationsbezogen eine Echokardiographie erhielten, als mögliche Teilnehmer der Studie erfasst. Als Voraussetzung galt, dass eine Doppelmessung durch den ersten Untersucher und eine zeitnahe Kontrollmessung durch einen zweiten Untersucher möglich war. Ein probandenbezogener Studienabbruch aufgrund unerwarteter Ereignisse, wie starke Unruhe, ersichtliche Schmerzen oder vermutliches Unwohlsein, trat nicht ein.



## 4.2 Datenerhebung und Dokumentation

Die Basisdaten eines jeden Probanden wurden aus den Datenerfassungssystemen Neodat bzw. COPRA entnommen.

Nach vorheriger Absprache mit dem Pflegepersonal wurden die NG nacheinander von 2 Ärzten in einem zeitlich begrenzten Rahmen von insgesamt maximal zehn Minuten am Bettplatz untersucht. Der eigentliche Untersuchungsablauf durch beide Untersucher nahm dabei maximal 5 Minuten in Anspruch. Die Studie erweiterte die Standarduntersuchung auf eine doppelte Messung (1. Und 2. STI-Messung) durch den ersten Untersucher und eine weitere verblindete Messung (3. STI-Messung) durch einen zweiten Untersucher. Die gesamte Untersuchungszeit verlängerte sich also auf das 3-Fache der „normalen“ Zeit (je Messung ca. 60 bis 120 s). Dabei wurde strikt auf gleichbleibende Untersuchungsvoraussetzungen geachtet (gleiche Lichtverhältnisse, keine Störfaktoren wie Lärm etc.).

Die Dokumentation der für unsere Studie ermittelten Daten fand auf einem zuvor ausgearbeiteten Bogen statt. Dieser wurde vor Untersuchungsbeginn mit einem Patientenaufkleber mit dem Namen/Nachnamen, der Patientennummer und dem Geburtsdatum des Kindes versehen.

Im Anschluss erfolgte die Durchführung der echokardiographischen Ermittlung der STIs nach Literaturvorgaben (siehe 4.3). Dabei erhob eingangs der erste Arzt sowohl die Dauer der lv (LPEP, LVET), als auch der rv Zeitintervalle (RPEP, RVET), sowie die HF. Dies geschah anhand zwei aufeinanderfolgender Herzzyklen, um später die intraindividuelle Abweichung analysieren zu können. Im Nachgang erfolgte die erneute Messung selbiger Werte auf die gleiche Weise anhand eines Herzzyklus durch den zweiten Untersucher. Alle Werte wurden auf dem Dokumentationsbogen festgehalten.

Die weiteren, während der echokardiographischen STI-Ermittlung erfassten, relevanten Informationen wurden ebenfalls auf dem Dokumentationsbogen notiert, welcher vom Studienleiter verwahrt wurde.

Dazu zählten das aktuelle Alter des Kindes zum Zeitpunkt der Untersuchung, sowie das aktuelle Gewicht, ob das Kind unter dem Einfluss von Katecholaminen (KA) stand, ob während des Herzechos ein offener PDA festgestellt wurde und wie agitiert die Probanden zum Untersuchungszeitpunkt waren (Darstellung auf VAS).

Zur Abschätzung der Vergleichbarkeit der Untersuchungsbedingungen hinsichtlich Agitiertheit/Unruhe wurden zwei Angaben gemacht: zum einen wurde die Vigilanz des Kindes mit der VAS bestimmt, zum anderen wurde die HF des Kindes genau dokumentiert. Die VAS ist eine subjektive Skala zur Einschätzung von „Zuständen“ und wird häufig zur Erfassung von Schmerzen eingesetzt, kann aber auch zur Beurteilung anderer Charakteristika benutzt werden. Die Vigilanz wurde durch eine Markierung auf einer nicht unterteilten geraden Linie festgehalten. Die zwei Enden der Linie charakterisieren die jeweiligen Extreme von links (ruhiges Kind) bis rechts (maximal agitiertes Kind). Die subjektive, rein visuelle Einschätzung wurde anschließend im Rahmen der statistischen Auswertung als prozentualer Wert quantifiziert. Hierfür wurde die Länge bis zur Position der Markierung gemessen und das Verhältnis zur Gesamtlänge der Linie gebildet. Die Angabe der Ergebnisse erfolgte dann in Prozent mit naturgemäß möglichen Werten von 0-100%. Erstmals wurde die VAS 1921 von Hayes & Patterson publiziert.<sup>39</sup>

Wir definierten eine Abweichung von maximal 10% der prozentualen Werte als „gleiche“ Vigilanz. Für die Abweichung der HF wurden bis zu 5% als optimale Voraussetzung hinsichtlich der Wiederholbarkeit der Methode festgelegt.

Der Leiter der Studie behielt außerdem eine entsprechende Liste sämtlicher Untersucher mit den jeweils zugeordneten Nummern hinsichtlich der Expertise in der Echokardiographie (zum Zeitpunkt der Untersuchung) ein. Die Untersucher wurden nicht im Voraus festgelegt, sämtliche eingearbeitete Mitarbeiter durften an der Datenerhebung teilnehmen, um die Studie möglichst alltagsgetreu zu gestalten. Obwohl alle Untersucher gut instruiert und vertraut mit der Messung der STIs waren, wurden zwei Gruppen von Untersuchern hinsichtlich ihres Erfahrungslevels (EL) gebildet (<1 Jahr Erfahrung = EL 1, >1 Jahr Erfahrung = EL 2).

Nachträglich wurden weitere für unsere Studie wichtige Daten, wie die absolvierten Schwangerschaftswochen (SSW) bei Geburt des Kindes, die Entwicklungsgrößen GG und Trophe, der 5-Minuten-APGAR und der Nabelschnur-pH aus dem Neodat/COPRA ermittelt und auf dem Dokumentationsbogen eingetragen.

Das generierte Ultraschallbild wurde vom jeweiligen Untersucher mit dem Namen des Patienten versehen, gespeichert und nach der Untersuchung ausgedruckt. Alle erhobenen STI-Werte befinden sich auf dem Bild. Selbiges wurde in einem eigens dafür vorgesehenen Ordner abgelegt, zusammen mit dem Patientenbogen. Der Ordner wurde durch den Studienleiter verwahrt.

**Anmerkung:**

*Bei einigen statistischen Auswertungen konnten nicht alle, normalerweise dafür in Frage kommenden Untersuchungen, einbezogen werden. Es kommt zu Abweichungen aufgrund fehlender Einzelmessungen einer der beiden Untersucher oder mangelnder Dokumentation. Die Anzahl der in die Analysen einfließenden Untersuchungen ist daher individuell dokumentiert. Insgesamt sind bei 5 Untersuchungen lv STIs und bei 9 Untersuchungen rv STIs unvollständig dokumentiert. Bei fehlenden Daten bzgl. der STIs handelt es sich in der Regel um die 2. Messung der STIs durch den ersten Untersucher. In 2 Messreihen konnten die rv STIs aufgrund einer ausgeprägten Lungenüberblähung beim Probanden nicht erhoben werden. Bei 10 Messungen wurde die VAS durch den 2. Untersucher nicht festgehalten. Teilweise wurden Angaben für die Subgruppenanalyse (Katecholamingabe, Sauerstofftherapie etc.) nicht dokumentiert und waren nachträglich nicht recherchierbar.*

#### 4.3 Theorie und Durchführung der Echokardiographie zur STI-Ermittlung

Zur Ermittlung der STIs wurde ein ProSound Alpha 7 (Aloka, Hitachi-Europe, Zug, Switzerland) Ultraschallgerät verwendet. Dabei konnten, bei Vorhandensein eines integrierten EKG-Gerätes, die patienteneigenen Elektroden, die außerhalb der Untersuchung dem Monitoring dienten, an das Gerätekabel umgesteckt werden. Zusätzliche Elektroden mussten demnach nicht angebracht werden. Die EKG-Ableitung erfolgte zeitgleich während der gesamten Untersuchung.

Mit Hilfe eines 5 MHz Sektorschallkopfes wurde bei den Probanden im Zusammenhang mit der postpartalen Erstuntersuchung auf kongenitale Herzfehler, einen PDA o.Ä., die Analyse der STIs durchgeführt. Das NG lag während der Realisierung des Herzechos in Rücken- oder Linksseitenlage. Um die parasternal lange Achse darzustellen, wurde der Schallkopf im 2./3. oder 4. Interkostalraum links parasternal auf dem Thorax des Kindes platziert und ein zweidimensionales Ultraschallbild im B-Mode aufgezeichnet. Die kurze Achse wurde nach einer Drehung des Schallkopfes um 90° im Uhrzeigersinn eingestellt. Man nutzt diese Achse nicht nur zur Detektion eines PDA (siehe Abbildung 4), sondern auch, um mit Hilfe des Pulsed wave (PW) - Doppler die linksseitigen STIs, LPEP und LVET zu erheben. Nach optimaler Darstellung der Aorta und ihrer Taschenklappen, wurde der Cursor zentral in der Klappenebene positioniert und ein Flussprofil abgeleitet. Anschließend wurden senkrechte Hilfslinien über das Flusssignal der Aorta gelegt, eine zur Markierung der Klappenöffnung und eine zur Kennzeichnung des Klappenschlusses, dem Ende der Austreibungsphase. Danach konnten die LVETs automatisiert durch das Echogerät in Millisekunden berechnet werden. Die LPEP umfasst das Intervall zwischen der Q-Zacke im parallel abgeleiteten EKG und der Aortenklappenöffnung (siehe Abbildung 2).

Analog erfolgt die Messung der rv STIs, jedoch über der Pulmonalklappe, die ebenfalls in der parasternalen kurzen Achse dargestellt wird (siehe Abbildung 3).

Die Summe aus PEP und VET bezeichnet die gesamte elektromechanische Systole und wird für die linksseitigen STIs auch als  $QA_2$  bezeichnet, die Summe aus RPEP und RVET wiederum  $QP_2$ .<sup>15</sup>

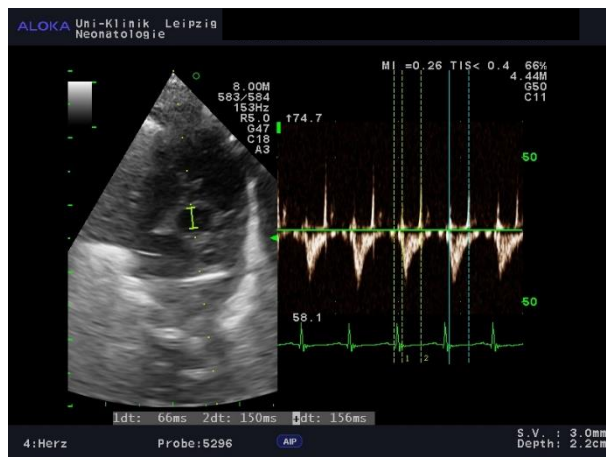


Abbildung 2 Messung STI über Aortenklappe

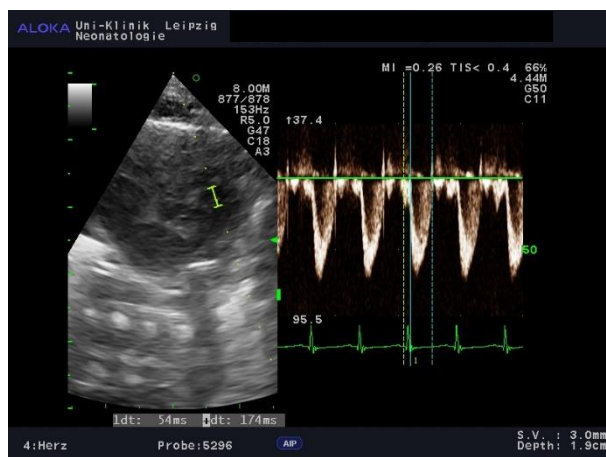


Abbildung 3 Messung STI über Pulmonalklappe

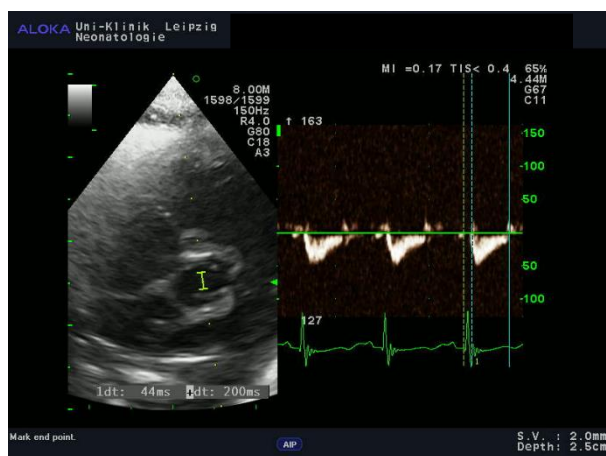


Abbildung 4 Messung STI über Aortenklappe bei bestehendem PDA

## 4.4 Grundlagen zur Statistik

### 4.4.1 Allgemeine Statistik

Wir haben zunächst die Rohdaten der STIs anhand der Dokumentationsbögen zusammengetragen und hinsichtlich der Normalverteilung untersucht. Nicht alle STI-Werte waren nach dem Shapiro-Wilk-Test normalverteilt, gemäß des zentralen Grenzwertsatzes haben wir bei einer Stichprobengröße  $n > 30$  dennoch zu allen Daten die Mittelwerte und die entsprechende Standardabweichung dokumentiert.

### 4.4.2 Bland et Altmann Plots

Zur Veranschaulichung der Schwankungsbreite der Differenzen zwischen den Werten intraindividuell/interindividuell wiederholter Messungen der STIs, wurden in unserer Studie Diagramme nach dem Prinzip von Bland et Altmann verwendet. Diese können gleichzeitig mögliche systematische Messfehler aufzeigen. Die Punktdiagramme tragen die Differenzen der jeweilig wiederholten Messungen der einzelnen Untersuchungen gegenüber dem Mittelwert der 2 Messwerte einer jeden Untersuchung auf. Gleichzeitig wird aus der Summe aller Differenzen erneut der Mittelwert gebildet und im plot entsprechend als waagerechte Linie zur x-Achse markiert. Ebenso wird die doppelte Standardabweichung (SD), sowohl ins Positive, als auch ins Negative eingetragen. Der Bereich zwischen der doppelten SD ins Positive und ins Negative gibt das sogenannte 95%-Konfidenzintervall (95%-KI) an, in dem 95% aller Differenzen der einzelnen wiederholten Messungen zu erwarten sind. Messunterschiede innerhalb des entsprechenden Intervalls gelten als klinisch nicht relevant (limits of agreement).<sup>40</sup> Die Bland et Altmann plots wurden mit Hilfe des Statistikprogrammes SPSS, Version 22, generiert.

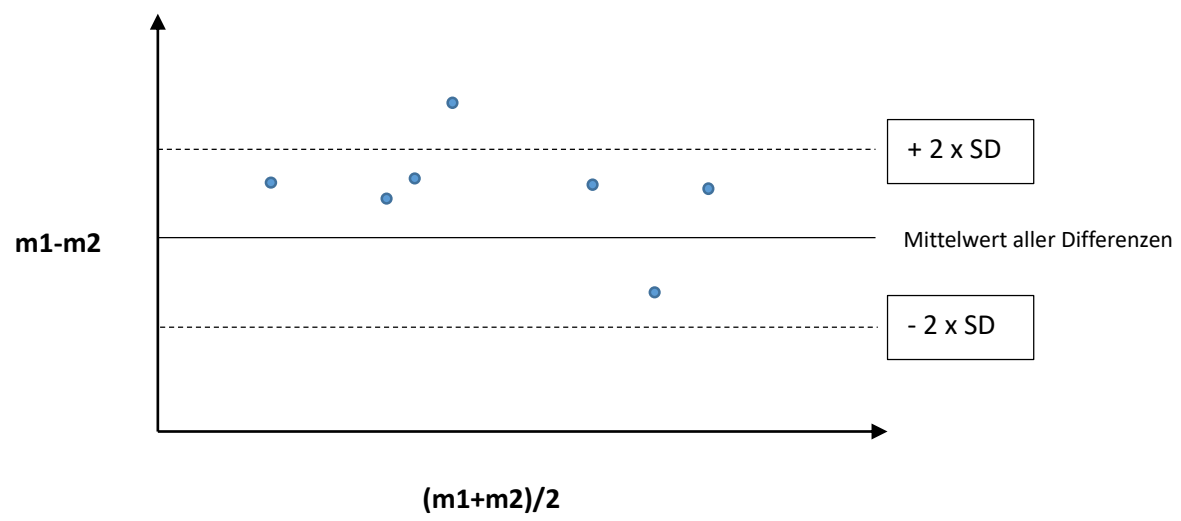


Abbildung 5 Beispielhafte Darstellung eines Bland et Altman Diagrammes

#### 4.4.3 Repeatability Coefficient und Repeatability Index

Die erste Auswertung der Daten erfolgte anhand der statistisch angewandten Methode der Studie von Schwarz et al.<sup>31</sup> Der repeatability coefficient (RC) ist ein Maß, um die Präzision wiederholter Messungen zu beschreiben. Man errechnet jeweils die Differenz der zwei Messungen im Rahmen einer Untersuchung und bildet am Ende den Mittelwert der Differenzen sämtlicher Untersuchungen. Der RC stellt die 2-fache SD gegenüber diesem Mittelwert der Differenzen dar.

$$(U1.M1-U1.M2)^2 + (U2.M1-U2.M2)^2.... + (Un.M1-Un.M2)^2 / n = X$$

- U = Untersuchung
- M = Messung
- n = Anzahl der Untersuchungen
- $\sqrt{X}$  = SD of the differences
- RC = 2 x SD of the differences

Dabei wird erwartet, dass 95% der errechneten Differenzen kleiner sind als die doppelte Standardabweichung. Anders formuliert, der RC beschreibt den Erwartungswert, unter dem mit einer Wahrscheinlichkeit von 95% die absolute Differenz zwischen zwei Ergebnissen einer wiederholten STI-Analyse liegt. Somit kann anhand des RC die Zuverlässigkeit von Mehrfachmessungen aufgezeigt werden.<sup>40</sup>

Der Repeatability index (RI) ist das Ergebnis der Division des RC durch den Mittelwert aller Messungen.<sup>31</sup> Er stellt somit das Verhältnis der doppelten Standardabweichung des Mittelwertes der Differenzen sämtlicher Messungen gegenüber dem Mittelwert der Ergebnisse aller Messungen dar.

$$RI = RC / ([M1+M2....Mn] / n)$$

Der RC zeigt im Vergleich repetitiver Messungen also die Streuung der absoluten Differenzen auf, ohne Bezug zur Größe des eigentlichen Messwertes zu nehmen. Der RI normiert die Streuung der Differenzen der Messwerte auf den Mittelwert aller Messwerte. Somit lässt er eine prozentuale Beurteilung der Übereinstimmung wiederholter Messungen.



Nach Rücksprache mit den Mitarbeitern des IMISE innerhalb der statistischen Beratung wurde, in Ermangelung eines offiziellen Grenzwertes, ein Grenzwert von 0,1 (also 10%) als akzeptabler Wert für den RI festgelegt.

Die Standardabweichung der Differenzen und der Mittelwert der Messungen wurden mit Hilfe des Statistikprogrammes SPSS berechnet.

Wir untersuchten die akzeptablen RIs im Anschluss hinsichtlich der zuvor festgelegten möglichen Einflussfaktoren (Erfahrungslevel, Geschlecht, SST, GG, APGAR, O<sub>2</sub>-Therapie, Beatmung, Katecholamintherapie, PDA). Dabei betrachteten wir vorrangig die Maxima/Minima. Weiterhin analysierten wir sämtliche RIs hinsichtlich deutlicher Tendenzen innerhalb der einzelnen Kategorien mit ebenfalls zuvor definierten Subgruppen:

- Erfahrungslevel                      EL1/1, EL1/2, EL2/1, EL2/2
- Geschlecht                            männlich/weiblich
- SST                                      <259, <196
- GG                                       <1500g, ≥1500g
- APGAR                                  <5, 5-7, ≥8
- O<sub>2</sub>-Therapie                            ja/nein
- Beatmung                                ja/nein
- Katecholamintherapie                ja/nein
- PDA                                        ja/nein

Für alle genannten Untersuchungen wurden die Werte anschließend nach Abweichungen der Vigilanz und HF ( $VAS1-VAS2 \leq 10\%$ ;  $(HF1-HF2) \leq 0,05 \cdot HF1$ ) gefiltert und sämtliche Betrachtungen unter Beachtung dieser Bedingungen erneut durchgeführt.

#### 4.4.4 Cohens Kappa coefficient

Aufgrund mangelnder qualitativer Beurteilung der Daten anhand des RC oder des RI hinsichtlich der klinischen Relevanz der Wiederholbarkeit der Messungen haben wir uns entschieden, den Datensatz erneut mit Hilfe des Kappa-Koeffizienten nach Cohen auszuwerten. Dies stellt eine qualitative Analyse der Datenabweichung dar.

Die statistische Methode kann dabei sowohl bei der Bewertung intraindividuelle, als auch interindividuelle Messungen angewendet werden. Das Maß der Übereinstimmung wird dann als intraindividuelle Reliabilität (ein Untersucher, 2 Messungen) bzw. interindividuelle Reliabilität (2 Untersucher, jeweils 1 Messung) bezeichnet.

Der Koeffizient beschreibt das Verhältnis zwischen der zufällig zu erwartenden Übereinstimmung  $p_e$  und der tatsächlich beobachteten Übereinstimmung  $p_o$ .

Ist Kappa gleich Null, bedeutet dies, dass die tatsächliche Übereinstimmung denselben Wert annimmt, wie die erwartete zufällige Übereinstimmung. Je höher der Kappa-Koeffizient ist, desto stärker ist also die tatsächliche Reliabilität.

$$K = (p_o - p_e) / (1 - p_e)$$

Einzelheiten zur mathematischen Berechnung sind der Quelle zu entnehmen <sup>41</sup>

Die ermittelten STIs wurden qualitativ in 3 Kategorien eingeteilt:

1 = Wert liegt oberhalb des Normbereiches

2 = Wert liegt innerhalb des Normbereiches

3 = Wert liegt unterhalb des Normbereiches

Dabei wurden die PEPs mit Hilfe der auf das Gestationsalter und den LT abgestimmten Normbereiche ( $</ \geq 5$ . LT) nach Halliday et al. eingeteilt.<sup>42</sup> Die Limits für die individuellen Grenzwerte der LVETs wurden unter Einsatz der Formel

$$240 - \frac{1}{2} Hf \pm 0,05 \times (240 - \frac{1}{2} Hf)$$

erstellt. Diese Formel entspricht der internen Leitlinie der Abteilung Neonatologie der Klinik und Poliklinik für Kinder- und Jugendmedizin Universität Leipzig und wurde in Voruntersuchungen als Standard für die Frequenzkorrektur ermittelt.

Die Bewertung der erhobenen Werte für die RVETs wurde ebenfalls nach dem Normbereich nach Halliday et al durchgeführt.<sup>42</sup>

Die Auswertung erfolgte mit Hilfe des Statistikprogrammes SPSS. Das Maß der Übereinstimmung unserer Untersuchungen wurde anhand der ermittelten Kappa Koeffizienten nach den von Landis et al angegebenen Normen interpretiert.<sup>43</sup>

*Tabelle 1 Interpretation Kappa Werte nach Landis und Koch*

Kappa Wert	Übereinstimmung
0–.20	gering
.21–.4	ausreichend
.41–.6	moderat
.61–.8	beachtlich
.81–1.0	(fast) vollständig

Es gibt unterschiedliche Normwerte der STIs abhängig vom Gestations-/Lebensalter, weshalb wir, aufgrund ihrer deutlichen Überzahl in unserer Kohorte, ausschließlich FG bei der Auswertung der Daten nach Cohens Kappa einbezogen. Da die PEPs mit Hilfe der auf den LT abgestimmten Normbereiche ( $</\geq 5$ . LT) nach Halliday et al. bewertet wurden, haben wir die Daten bei der Kappa-Analyse zusätzlich in entsprechende Subgruppen einer Kategorie „Lebenstag“ unterteilt und untersucht.<sup>42</sup>

Wir untersuchten die Werte wieder hinsichtlich der zuvor festgelegten möglichen Einflussfaktoren (Erfahrungslevel, Geschlecht, SST, GG, APGAR, O<sub>2</sub>-Therapie, Beatmung, Katecholamintherapie, PDA). Erneut achteten wir auf ersichtliche Tendenzen innerhalb der einzelnen Kategorien (Subgruppen entsprechend 4.4.3).

Für alle genannten Untersuchungen wurden die Werte anschließend nach Abweichungen der Vigilanz/HF ( $VAS1-VAS2 \leq 10\%$  /  $(HF1-HF2) \leq 0,05 \cdot HF1$ ) gefiltert und sämtliche Betrachtungen unter Beachtung dieser Bedingungen erneut durchgeführt.

*Tabelle 2 Beispielhafte Darstellung einer generierten Kreuztabelle mit SPSS im Rahmen der Datenauswertung anhand des Kappa-Koeffizienten nach Cohen*

			Normbereich LVET3			Gesamtsumme
			1	2	3	
Normbereich LVET1	1	Anzahl	7	10	1	18
		Erwartete Anzahl	3,4	9,7	4,9	18,0
	2	Anzahl	8	28	5	41
		Erwartete Anzahl	7,7	22,2	11,1	41,0
	3	Anzahl	1	8	17	26
		Erwartete Anzahl	4,9	14,1	7,0	26,0
Gesamtsumme		Anzahl	16	46	23	85
		Erwartete Anzahl	16,0	46,0	23,0	85,0

1 = Wert liegt oberhalb des Normbereiches  
2 = Wert liegt innerhalb des Normbereiches  
3 = Wert liegt unterhalb des Normbereiches

LVET jeweils 1. Messung = LVET1 (von erstem Untersucher durchgeführt) gegenüber 3. Messung = LVET3 (vom zweiten Untersucher durchgeführt)

### Erklärung Tabelle 2:

Die Tabelle stellt die erwartete Anzahl der zufällig übereinstimmenden Messergebnisse der tatsächlich beobachteten Anzahl der Übereinstimmungen gegenüber.

Gezeigt sind die individuell durch beide Untersucher gemessenen LVET-Werte (LVET1= 1. Messung 1. Untersucher; LVET3= Messung 2. Untersucher). Dabei wurden die Messwerte qualitativ verglichen und nach den Kriterien „oberhalb des Normbereichs“ (1), „innerhalb des Normbereichs“ (2) und „unterhalb des Normbereichs“ (3) einsortiert. Bei den dickgedruckten Ergebnissen sind die Messungen der beiden Untersucher nach diesen Kriterien gleich.

Eine höhere Anzahl der tatsächlichen Übereinstimmungen gegenüber den statistisch erwarteten (zufälligen) Übereinstimmungen ist prinzipiell wünschenswert.

Die konkreten Werte in der Tabelle bedeuten: Übereinstimmung der Werte zwischen LVET1 und 3 „oberhalb des Normbereichs“ bei 7 Messungen (erwartet 3,4), „innerhalb des Normbereichs“ bei 28 Messungen (erwartet 22,2) und „unterhalb des Normbereichs“ bei 17 Messungen (erwartet 7). Somit ergibt sich eine (qualitative) Übereinstimmung bei 52 von 85 Messungen. Statistisch zufällig zu erwarten wären nur 32,6 Übereinstimmungen.

Stimmen die Ergebnisse der Untersucher nicht überein (z.B. LVET1 =1 „oberhalb des Normbereichs“; LVET3 =3 „unterhalb des Normbereichs“), ist eine niedrigere Anzahl der Messungen (in dem Fall n=1) gegenüber der statistisch erwarteten Anzahl der betroffenen Messungen (in dem Fall n=4,9) anzustreben (siehe Tabelle 2, kursiv gedruckt).

*Tabelle 3 zugehörig zu Tab. 2, berechneter Kappa-Wert nach  $K = (p_o - p_e) / (1 - p_e)$*

<b>Symmetrische Maße</b>				
	Wert	Asymp. Standardfehler <sup>a</sup>	Näherungsweis e A <sup>b</sup>	Näherungsweis e Sig.
Maß für die Übereinstimmung	Kappa	,370	,085	4,702
Anzahl der gültigen Fälle	85			,000

a. Die Nullhypothese wird nicht vorausgesetzt.

b. Unter Annahme der Nullhypothese wird der asymptotische Standardfehler verwendet.

## 5. Ergebnisse

### 5.1 Basisdaten

98 Untersuchungen mit jeweils 3 Messungen an insgesamt 75 Kindern wurden durchgeführt. Die Patienten befanden sich zum Zeitpunkt der Untersuchung entweder auf der NG Intensiv- oder Normalstation. 43 (57,3%) der NG waren männlich und 32 (42,7%) weiblich. Es zeigte sich, dass 65 Patienten (86,7%) unserer Kohorte FG waren, davon 29 (38,7%) FG mit extremer Unreife. 47 (62,7%) Kinder kamen mit einem GG kleiner als 1500g zur Welt. Das durchschnittliche GG betrug 1455g. Die Kinder erreichten im Schnitt die 39. Gewichts-Perzentile.

Nicht alle Messungen fanden am 1. LT statt, weshalb das Gewicht zum Zeitpunkt der Datenerhebung (im Durchschnitt ca. 11. LT) bei 1700g lag. Über die Hälfte der STI-Bestimmungen erfolgten vor dem 5. LT (ca. 55%).

In 63/98 (64,3%) der Messungen standen die NG unter dem Einfluss einer KA-Therapie (Dobutamin, Adrenalin, Dopamin), in 76,5% der Fälle hatten die Patienten einen erhöhten Sauerstoffbedarf. 43,9% der Untersuchungen fanden unter Beatmung statt. Ein PDA konnte in ca. 44% der Betrachtungen nachgewiesen werden.

Die Daten hinsichtlich der vergangenen Schwangerschaftstage (SST) bis zum Geburtszeitpunkt, der Trophe, des GG, des APGAR-Wertes, des Nabelarterien-pHs (NA-pH) und des Gewichts zum Zeitpunkt der Untersuchung waren nicht normalverteilt.

Tabelle 4 Übersicht Basisdaten

<b>Schwangerschaftstage (d)</b>	<b>209 ± 34</b> <hr/> <b>29/75 &lt;196 SST (38,7%)</b> <b>48/75 &lt;224 SST (64,0%)</b> <b>65/75 &lt;259 SST (86,7%)</b>	<b>Lebenstag (d)</b>	<b>11 ± 15</b> <hr/> <b>54/98 &lt; 5. LT (55,1%)</b> <b>44/98 ≥ 5. LT (44,9%)</b>
<b>Trophe (%)</b>	<b>39 ± 27</b>	<b>APGAR (5.Minute)</b>	<b>7 ± 1,7</b> <hr/> <b>6/75 APGAR &lt;5 (8,0%)</b> <b>37/75 APGAR 5-7 (50,0%)</b> <b>32/75 APGAR ≥8 (42,7%)</b>
<b>Geburtsgewicht (g)</b>	<b>1455 ± 960</b>	<b>Gewicht (g)</b>	<b>1700 ± 1010</b>
<b>NA-pH</b>	<b>7,3 ± 0,4</b>	<b>PDA</b>	<b>43/98 ja (43,9%)</b> <b>54/98 nein (55,1%)</b> <b>1 unbekannt</b>

Dargestellt sind Daten bzgl. der gesamten Kohorte (75 Neugeborene), die Werte bzgl. der Eigenschaften Lebenstag/Gewicht /PDA, beziehen sich auf den Status der 75 Neugeborenen am Tag der Untersuchung, somit sind alle 98 Messungen einbezogen, Abkürzungen: NA-pH = Nabelarterien-pH, PDA = persistierender Ductus arteriosus. Werte angegeben als Mittelwerte ± Standardabweichung

## 5.2 Übersicht der Ergebnisse der Messungen der STIs

### 5.2.1 Untersuchungsvoraussetzungen

Im Rahmen der Studie gingen wir davon aus, dass zur optimalen Beurteilung der Reproduzierbarkeit und Verlässlichkeit der Methode, die Untersuchungsvoraussetzungen hinsichtlich der kindlichen Vigilanz und KL-Stabilität (VAS und HF) möglichst gleich sein sollten. Gleichzeitig sollte das EL der Untersucher (Level 1: <1 Jahr Erfahrung oder 2 :>1 Jahr Erfahrung) berücksichtigt werden. Diese Parameter wurden deshalb zuerst analysiert.

Der Erstuntersucher hatte in ca. 76,5% der Fälle EL 2. Der zweite Untersucher besaß in ca. 54% der Untersuchungen EL2.

Bei Analyse der HF und VAS als wichtige, die Hämodynamik beeinflussende Parameter, zeigte sich, dass die Zirkulation und die Vigilanz der Kinder, in den Mittelwerten, gleichblieben (siehe Tabelle 5). Die HF zu den unterschiedlichen Messzeitpunkten waren jeweils normalverteilt. Die Werte der VAS, die während der Messung des 2. Untersuchers erhoben wurden, waren ebenfalls normalverteilt.

*Tabelle 5 Übersicht Untersuchungsvoraussetzungen*

	UNTERSUCHER 1		UNTERSUCHER 2	
<b>ERFAHRUNGSLEVEL</b> <b>1 = &lt;1 JAHR ERFAHRUNG</b> <b>2 = &gt;1 JAHR ERFAHRUNG</b>	1	23/98 (23,5%)	1	44/97 (45,4%)
	2	75/98 (76,5%)	2	53/97 (54,1%)
<b>HERZFREQUENZ (1/MIN)</b>	155 ± 20		154 ± 20	
<b>VAS (%)</b>	30 ± 20		30 ± 20	

Dargestellt sind die Daten hinsichtlich der Untersuchungsvoraussetzungen für Untersucher 1 und Untersucher 2, Abkürzungen: VAS- visuelle Analogskala, Werte angegeben als Häufigkeiten (%), Mittelwerte ± Standardabweichung

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2, es wurde nur in 97 Messungen das EL des 2. Untersuchers dokumentiert



In einer detaillierteren Betrachtung der Untersuchungsvoraussetzungen konnten wir feststellen, dass sich die Untersucherexpertise in 43 % der Untersuchungen nicht unterschied (davon 37,1% EL 2/2). Dabei war die Anzahl der Untersuchungen, in denen beide Ärzte EL 1 zuzuordnen waren, zu gering (6,2%), um sie weiterführend zu untersuchen.

In ca. 39% der Datenerhebungen lag die Konstellation EL 2/1 vor. Bei ca. 18% der Messungen wies der erste Arzt hinsichtlich der Echokardiografie und der Erhebung von STIs weniger Erfahrung auf (EL 1/2).

Rund 80% der Messungen fanden, hinsichtlich der Abweichung der HF zu den einzelnen Messpunkten, unter idealen Bedingungen statt (siehe 4.2). Bei der Abweichung der Vigilanz zwischen den wiederholten Messungen waren ca. 64 % der Untersuchungen unter gleichen Bedingungen erfolgt (siehe Tabelle 6).

*Tabelle 6 Differenzen der Parameter zu den Untersuchungsvoraussetzungen*

<b>HF1-HF2 (1/min)</b>		1.24 ± 7.25
		<b>(HF1-HF2) ≤ 0,05*HF1: 78/98 (79,6%)</b>
<b>VAS1-VAS2 (%)</b>		9 ± 8
		<b>VAS1-VAS2 ≤ 10%: 56/88 (63,6%)</b>
<b>EL1-EL2 = 0</b>	(Expertise beider Untersucher gleich)	42/97 (43,3%)
<b>EL1 = 2 / EL2 = 2</b>	(beide > 1 Jahr Erfahrung)	36/97 (37,1%)
<b>EL1 = 1 / EL2 = 1</b>	(beide < 1 Jahr Erfahrung)	6/97 (6,2%)
<b>EL1 = 2 / EL2 = 1</b>	(Expertise 2. Untersucher geringer)	38/97 (39,2%)
<b>EL1 = 1 / EL2 = 2</b>	(Expertise 2. Untersucher höher)	17/97 (17,5%)

Dargestellt sind die Daten hinsichtlich der Untersuchungsvoraussetzungen für Untersucher 1 und Untersucher 2, Abkürzungen: HF = Herzfrequenz, HF1 = Herzfrequenz Messung 1. Untersucher; HF2 = Herzfrequenz Messung 2. Untersucher; VAS1 = visuelle Analogskala Messung 1. Untersucher, VAS2 = visuelle Analogskala Messung 2. Untersucher; EL1 = Erfahrungslevel 1. Untersucher, EL2 = Erfahrungslevel 2. Untersucher; Werte angegeben als Häufigkeiten (%), Werte angegeben als Mittelwerte ± Standardabweichung

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2, es wurde nur in 97 Messungen das EL des 2. Untersuchers dokumentiert, in 10 Fällen wurde die VAS durch den 2. Untersucher nicht dokumentiert

## 5.2.2 Übersicht zur durchschnittlichen Dauer der erhobenen STIs und zu den ermittelten WR

Die Darstellung der STIs demonstriert, dass die LPEPs durchschnittlich den RPEPs gleichen. Bei den VETs verhält es sich anders, die LVETs sind insgesamt kürzer als die RVETs.

*Tabelle 7 Mittelwerte der STIs aus allen 3 Messungen der Untersuchungen*

<b>LPEP</b> n=93 (ms)	55 ± 10	Prä-/Ejektionszeiten in Millisekunden, LPEP = linksventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, LWR = linksventrikuläre Weissler Ratio, RPEP = rechtsventrikuläre Präejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, RWR = rechtsventrikuläre Weissler Ratio; Werte angegeben als Mittelwerte ± Standardabweichung;  Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2, 5 lv STI-Messungen und 9 rv STI-Messungen unvollständig dokumentiert
<b>RPEP</b> n=89 (ms)	54 ± 12	
<b>LVET</b> n=93 (ms)	160 ± 20	
<b>RVET</b> n=89 (ms)	167 ± 20	
<b>LWR</b> n=93	0.35 ± 0.07	
<b>RWR</b> n=89	0.33 ± 0.08	

*Tabelle 8 Mittelwerte der STIs/WRs aller betrachteten Untergruppen der gesamten Kohorte*

	<b>LPEP</b>	<b>RPEP</b>	<b>LVET</b>	<b>RVET</b>	<b>LWR</b>	<b>RWR</b>
<b>MÄNNLICH</b>	54	54	159	165	0,35	0,33
<b>WEIBLICH</b>	56	54	162	168	0,35	0,33
<b>SST &lt;259</b>	54	52	158	165	0,35	0,32
<b>SST &lt;196</b>	53	49	152	158	0,35	0,32
<b>GG &lt;1500</b>	54	50	154	161	0,35	0,32
<b>GG ≥1500</b>	58	62	175	179	0,33	0,35
<b>O<sub>2</sub> JA</b>	54	53	159	165	0,34	0,33
<b>O<sub>2</sub> NEIN</b>	57	57	166	173	0,35	0,33
<b>KA JA</b>	54	55	164	172	0,34	0,33
<b>KA NEIN</b>	55	53	160	167	0,35	0,33
<b>PDA JA</b>	57	60	165	168	0,35	0,36
<b>PDA NEIN</b>	53	49	157	166	0,34	0,30
<b>APGAR &lt;5</b>	51	50	161	164	0,32	0,31
<b>APGAR 5-7</b>	55	53	159	168	0,35	0,32
<b>APGAR ≥8</b>	56	56	163	166	0,35	0,34
<b>BEATMUNG JA</b>	52	52	160	168	0,33	0,31
<b>BEATMUNG NEIN</b>	57	55	161	165	0,36	0,34
<b>Ø ALLE MESSUNGEN</b>	55	54	160	167	0,35	0,33

LPEP = linksventrikuläre Präejektionszeit; RPEP = rechtsventrikuläre Präejektionszeit; LVET = linksventrikuläre Ejektionszeit; RVET = rechtsventrikuläre Ejektionszeit; alle Zeitangaben in Millisekunden; LWR = Weissler ratio linksventrikuläre Zeitintervalle; RWR = Weissler ratio rechtsventrikuläre Zeitintervalle, SST = Schwangerschaftstage in Tagen (d); GG = Geburtsgewicht in Gramm (g); O<sub>2</sub> = Sauerstofftherapie (>21%); KA = Katecholamine; PDA = persistierender Ductus arteriosus botalli

### 5.3 Betrachtung der Daten mittels Bland et Altmann plots

Die Werte wurden intraindividuell (Messung 1 und 2 durchgeführt durch den ersten Untersucher) und interindividuell (Messung 1 durchgeführt durch den ersten Untersucher und Messung 3 durchgeführt durch den zweiten Untersucher) gegenübergestellt. Dabei wurden die Ergebnisse hinsichtlich des Vorliegens bereits bekannter Einflussfaktoren untersucht. Hierbei wurden insbesondere die „Ausreißer“  $> +2x\ SD$  /  $< -2x\ SD$  vom Mittelwert der Differenzen betrachtet und deren Charakteristika in Tabelle 29-30 im Anhang zusammengefasst.

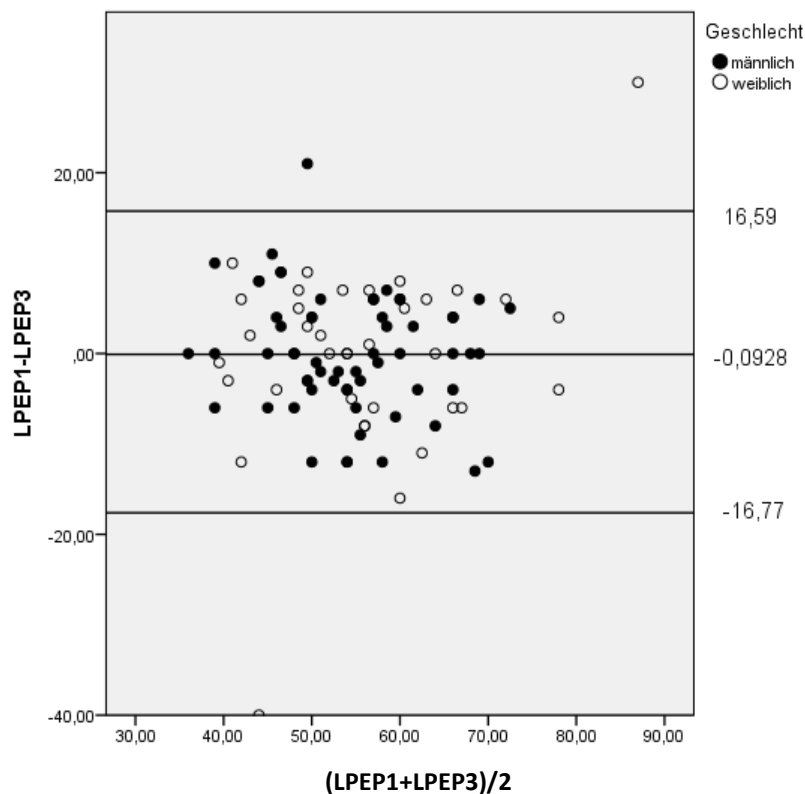
#### Linksventrikuläre STIs:

Bei der Betrachtung der lv STIs lagen 97% (91/94) der intraindividuellen Messunterschiede der PEP, bzw. 98% (92/94) der LVET im 95%-KI und sind somit nach dieser statistischen Methode als klinisch nicht relevant zu betrachten. Bezüglich der interindividuellen Abweichungen der Messwerte, befanden sich 97% (94/97) der Diskrepanzen der PEP-Werte und 96% (93/97) innerhalb des 95% KI.

Bei insgesamt 382 Messwertpaaren lagen somit nur 12 Werte außerhalb des 95%-KI. Diese 12 Werte wurden im Folgenden besonders analysiert, um systematische Abweichungen erkennen zu können. 10 der 12 Werte wurden männliche FG zugeordnet. Die Kinder wogen meist weniger als 1500g (7/12), 8 bekamen zusätzlichen Sauerstoff verabreicht. Die Mehrheit der NG besaß einen PDA (7/12), einige benötigten Katecholamine (5/12).

## Diagramm 1

*Beispiel für Bland et Altmann Diagramm: wiederholte Messung der LPEP durch unterschiedliche Untersucher- Betrachtung der Ausreißer hinsichtlich des Geschlechts*



y-Achse = Differenzen 1. Messung LPEP (durch Untersucher 1) ggü. 3. Messung LPEP (durch Untersucher 2); x-Achse = Mittelwert der erhobenen Messwerte durch Untersucher 1 und Untersucher 2; eingezeichnet wurden Markierungslinien des Mittelwerts der Differenzen  $\pm 2 \times SD$

### Rechtsventrikuläre STIs:

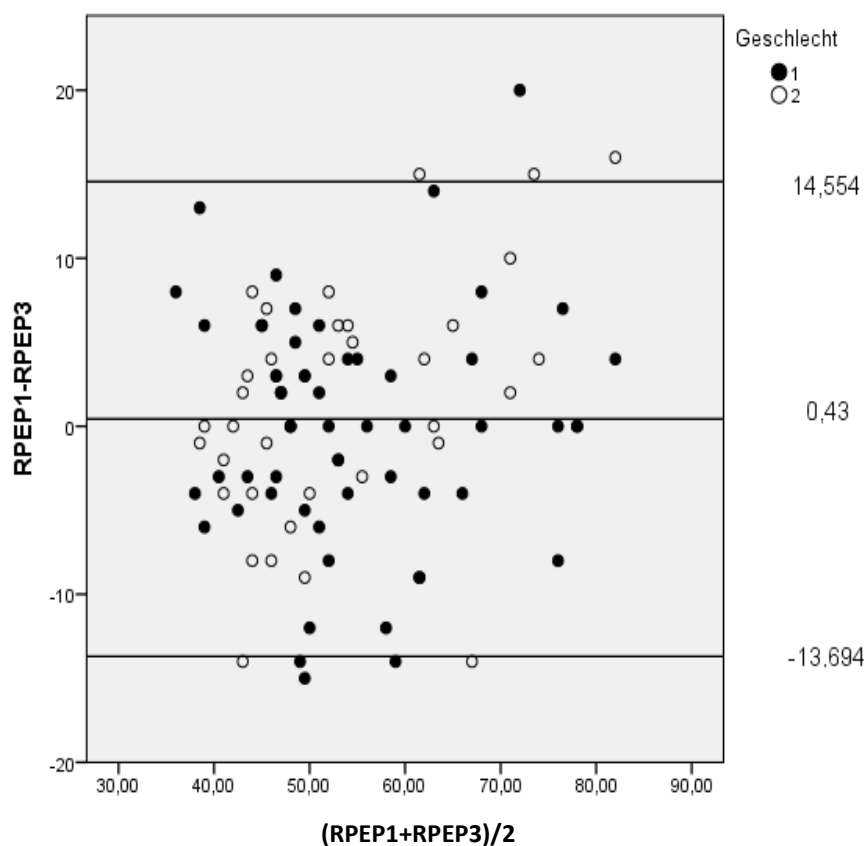
Die intraindividuellen Messunterschiede der RPEP lagen zu 98% (87/89) bzw. der RVET zu 94% (84/89) im 95%-Konfidenzintervall. Die Abweichungen bezüglich der interindividuellen Messwerte wiederum, befanden sich hinblickend auf die RPEP zu 91% (86/95) und bezogen auf die RVET zu 96% (91/95) innerhalb des 95%-Konfidenzintervalls. Alle Werte innerhalb des Intervalls sind anhand dieser Methode als klinisch nicht relevant einzuschätzen.

Bei der Analyse der rv STIs lagen also 20 Werte von 368 Messpaaren außerhalb des 95%-Konfidenzintervalls.

19/20 (95%) der Kinder sind FG. 12/20 (60%) der Kinder waren männlich und 75% der Messwerte gehörten zu Probanden mit weniger als 1500g GG. Bei 11 Ergebnissen wurde Sauerstoff über 21% insuffliert, 9 Kinder erhielten KA und In 8 Fällen lag ein PDA vor.

Diagramm 2

*Bland et Altmann Diagramm, wiederholte Messung der RPEP durch unterschiedliche Untersucher - Betrachtung der Ausreißer hinsichtlich des Geschlechts*



y-Achse = Differenzen 1. Messung RPEP (durch Untersucher 1) ggü. 3. Messung RPEP (durch Untersucher 2); x-Achse = Mittelwert der erhobenen Messwerte durch Untersucher 1 und Untersucher 2; eingezeichnet wurden der Mittelwert der Differenzen  $\pm 2 \times \text{SD}$

Anhand der Analyse mit Bland Altmann Plots zeigt sich, dass nur sehr wenige Werte außerhalb des Erwartungsbereiches lagen. Dies ist günstig, sagt aber noch zu wenig über die Vertrauenswürdigkeit der Daten aus. Im nächsten Schritt wurde explizit die Wiederholbarkeit untersucht.

## 5.4 Inter-/Intraindividuelle Wiederholbarkeit anhand des RC und dem RI

### Linksventrikuläre Messwerte (siehe Tabelle 9, Seite 33)

Der intraindividuelle RC für die LPEP lag bei 10,7, der RI bei 0,194 (erste Zeile). Der RI sagt damit aus, dass die Schwankungsbreite der Messwerte in einem Bereich von bis zu 19,4%, bezogen auf den durchschnittlichen Messwert, lag. Höher waren die Werte der interindividuellen LPEP Messungen (RC=16,7; RI=0,30). Mit anderen Worten: die Schwankungsbreite der Werte betrug dann bis zu 30% des durchschnittlichen Messwertes.

Der RC lag für intraindividuelle LVET-Messungen bei 12,2, der RI bei 0,076, was einer durchaus akzeptablen Schwankungsbreite von 7,6% entspricht. Die interindividuellen Messwerte lagen in einem höheren Bereich (RC=20,2; RI=0,126).

### Rechtsventrikuläre Messwerte (siehe Tabelle 9, Seite 33)

Der intraindividuelle RC für die RPEP lag bei 9,4, der RI bei 0,174 (erste Zeile). Die Schwankungsbreite der Messwerte betrug folglich, auf den durchschnittlichen Messwert bezogen, bis zu 17,4%. Bei interindividueller Messung lagen diese Werte in einem höheren Bereich (RC=14,1; RI=0,261).

Die Werte für die RVET waren bzgl. des RI abermals besser. Hier lag der RC für intraindividuelle Messungen bei 15,1, der RI bei ca. 0,09, was einer akzeptablen Schwankungsbreite von 9% entspricht. Für interindividuelle Messungen lagen diese Werte in einem höheren Bereich (RC=28,4; RI=0,17).

Im Vergleich zu den intraindividuellen Mehrfachmessungen, zeigen die interindividuell ermittelten Messwerte also in jedem Fall eine schlechtere Reproduzierbarkeit.

Die Ergebnisse zeigen außerdem, dass die Werte LVET und RVET intraindividuell ausreichend gut reproduzierbar sind (LVET1-2; RVET1-2).

Dies gilt mit gewissen Einschränkungen auch für interindividuelle Messungen. Bei den PEP-Messungen treten Schwankungen mit bis zu 30% des Mittelwertes der Messwerte auf, was an der Reproduzierbarkeit zweifeln lässt.

Es ergab sich die Frage, ob diese Reproduzierbarkeit abhängig von bestimmten Einflussfaktoren ist, weshalb die reproduzierbaren RIs hinsichtlich erkennbarer Tendenzen in verschiedenen Subgruppen analysiert wurden (siehe Schema 4.4.3).

#### Erfahrungslevel (siehe Tabelle 9)

Der RI für die LPEP lag unabhängig davon, ob die Untersucher dieselbe Erfahrung hatten, in einem Bereich um 23-25%. Auch die RI-Werte für die RPEP (ca.26-29%), die LVET (ca. 12-14%) und die RVET (ca. 15-18%) zeigten keine regelhafte Verbesserungstendenz hinsichtlich gleichbleibender Untersucherexpertise (EL2/2).

Die Untersuchungen durch zwei Untersucher mit dem EL 1 ist statistisch nicht verwertbar (n=4/n=6).

Folglich beeinflusst die Erfahrung der einzelnen Untersucher die Wiederholbarkeit der STI-Messungen nicht maßgeblich. Die niedrigsten RI-Werte wurden für die interindividuellen LVET-Messungen erreicht.

*Tabelle 9 Repeatability coefficient/Repeatability index der Differenzen aller Messwerte und gefiltert nach Erfahrungslevel*

		PEP Links		VET Links		PEP Rechts		VET Rechts	
		LPEP1-2	LPEP1-3	LVET1-2	LVET1-3	RPEP1-2	RPEP1-3	RVET1-2	RVET1-3
Alle	RC	10.7	16.7	12.2	20.2	9.4	14.1	15.1	28.4
	RI	<b>0,194</b> n=94	<b>0,303</b> n=97	<b>0,076</b> n=94	<b>0,126</b> n=97	<b>0,174</b> n=89	<b>0,261</b> n=95	<b>0,091</b> n=89	<b>0,170</b> n=95
EL 1/1	RC	/	16.0	/	7.9	/	11.6	/	52.9
	RI		0,330 n=6		0,052 n=6		0,208 n=4		0,290 n=4
EL 1/2	RC	/	27.9	/	20.1	/	15.6	/	25.9
	RI		0,234 n=17		0,121 n=17		0,285 n=17		0,148 n=17
EL 2/1	RC	/	13.0	/	21.3	/	13.5	/	28.6
	RI		0,243 n=38		0,137 n=38		0,256 n=38		0,178 n=38
EL 2/2	RC	/	13.8	/	19.8	/	14.5	/	26.3
	RI		0,248 n=36		0,121 n=36		0,265 n=36		0,156 n=36

RC = repeatability coefficient, RI = repeatability index, LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, ; LPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der LPEP; LPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der LPEP; entsprechend für LVET, RPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der RPEP; RPEP1-3 = Messung 1 Untersucher 1 – Messung 3 Untersucher 2 der RPEP; entsprechend für RVET, Maximum Rot, Minimum Grün, EL = Erfahrungslevel, EL 1/1 = beide Untersucher hatten Erfahrungslevel 1, EL 1/2= erster Untersucher hatte EL 1, zweiter Untersucher hatte EL 2 usw.

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2

Untersuchung der ausreichend reproduzierbaren intraindividuell erhobenen VETs  
(LVET1-2; RVET1-2)

Geburtsgewicht (siehe Tabelle 10, Seite 35)

Es zeigt sich, dass unter den akzeptablen Werten der LVET1-2, jene Untersuchungen der Kinder mit einem GG < 1500g den schlechtesten RI erzielten (RI=8,6%, rot markiert).

Wiesen die NG bereits bei Geburt ein Gewicht von 1500g oder mehr auf, war der ermittelte RI äußerst zufriedenstellend (RI=5,5% grün markiert).

Untersucht man parallel die normwertigen RIs der intraindividuellen RVET-Messungen (RVET1-2) hinsichtlich eines möglichen Einflusses des GG verhält es sich gleich, auch wenn die Werte insgesamt etwas höher lagen (RI-GG< 1500g=10,2%; RI-GG≥1500g=7%).

APGAR (siehe Tabelle 10, Seite 35)

Bei Betrachtung der akzeptablen RIs fällt weiterhin auf, dass die Werte der LVET1-2-Messungen bei Kindern mit einem 5-Minuten-APGAR von 5-7 (RI=8,6%) zu den schlechtesten/grenzwertig akzeptablen RIs gehören (rot markiert). Untersuchungen von Kindern mit einem APGAR ≥ 8 (RI=6,7%) erzielten akzeptable Werte.

Untersucht man parallel die normwertigen RIs der intraindividuellen RVET-Messungen (RVET1-2) hinsichtlich eines möglichen Einflusses des APGARs zeigt sich die gleiche Tendenz, auch wenn die Werte insgesamt etwas höher lagen (RI-APGAR 5-7=9,8%; rot markiert, RI-APGAR ≥ 8=8,3%,).

Die Werte der Subgruppe mit einem APGAR <5 sind bei zu niedriger Kohortenzahl (n=8/n=9/n=10) statistisch nicht verwertbar.



Tabelle 10 Repeatability coefficient/Repeatability index der Differenzen aller Messwerte gefiltert nach SST/GG/APGAR

		PEP Links		VET Links		PEP Rechts		VET Rechts	
		LPEP1-2	LPEP1-3	LVET1-2	LVET1-3	RPEP1-2	RPEP1-3	RVET1-2	RVET1-3
SST<259 (37.SSW)	RC	10.6	15.5	12.5	20.6	9.4	13.9	15.2	28.2
	RI	0,194 n=83	0,290 n=86	0,079 n=86	0,130 n=83	0,179 n=78	0,266 n=84	0,092 n=78	0,172 n=84
SST<196 (28.SSW)	RC	12.5	17.8	12.8	16.8	8.6	12.9	15.1	28.6
	RI	0,237 n=41	0,338 n=42	0,085 n=41	0,111 n=42	0,175 n=37	0,260 n=40	0,094 n=37	0,182 n=40
GG<1500 (g)	RC	10.9	15.9	13.2	18.9	8.7	14.8	16.4	30.3
	RI	0,202 n=64	0,296 n=66	0,086 n=64	0,123 n=66	0,173 n=59	0,295 n=64	0,102 n=64	0,189 n=59
GG≥1500 (g)	RC	10.3	17.8	9.5	22.8	10.9	12.4	12.5	23.0
	RI	0,181 n=30	0,308 n=31	0,055 n=30	0,130 n=31	0,176 n=30	0,203 n=31	0,070 n=30	0,128 n=31
APGAR <5	RC	10.7	21.9	9.8	25.6	8.6	16.5	11	40.8
	RI	0,210 n=10	0,428 n=10	0,061 n=10	0,159 n=10	0,175 n=8	0,332 n=9	0,066 n=8	0,251 n=9
APGAR 5-7	RC	12.2	12.6	13.6	18.4	9.2	11.9	16.5	24.9
	RI	0,222 n=45	0,229 n=47	0,086 n=45	0,116 n=47	0,173 n=43	0,227 n=46	0,098 n=43	0,148 n=46
APGAR ≥8	RC	8.7	19.7	11	21.1	9.7	16.0	13.7	28.7
	RI	0,154 n=39	0,350 n=40	0,067 n=39	0,130 n=40	0,172 n=38	0,283 n=40	0,083 n=38	0,173 n=40

RC = repeatability coefficient, RI = repeatability index, LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, ; LPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der LPEP; LPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der LPEP; entsprechend für LVET; RPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der RPEP; RPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der RPEP; entsprechend für RVET, Maximum Rot, Minimum Grün, SST = Schwangerschaftstage; GG = Geburtsgewicht

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2

### Katecholamine (siehe Tabelle 11, Seite 36)

Der RI der wiederholten LVET-Messungen durch denselben Untersucher an Kindern mit Katecholamintherapie zeigte äußerst zufriedenstellende Werte (RI=5,7%). Untersuchungen von Kindern ohne Katecholamingabe, lagen mit 8,5% etwas höher.

Untersucht man parallel die normwertigen RIs der intraindividuellen RVET-Messungen (RVET1-2) hinsichtlich eines möglichen Einflusses einer Katecholamintherapie, verhält es sich entgegengesetzt. Der RI der gemessenen RVET1-2 zeigt bei den Kindern, die KAs erhielten, eine noch akzeptable Wiederholbarkeit (RI=9,8% rot markiert). Untersuchungen von Kindern ohne Katecholamingabe (RI=8,3%) erzielten niedrigere Werte.

Tabelle 11 Repeatability coefficient/Repeatability index der Differenzen aller Messwerte gefiltert nach KA-Gabe/PDA

		PEP Links		VET Links		PEP Rechts		VET Rechts	
		LPEP1-2	LPEP1-3	LVET1-2	LVET1-3	RPEP1-2	RPEP1-3	RVET1-2	RVET1-3
KA ja	RC	9.4	15.6	9.4	22.4	9.5	12.3	16.9	33.8
	RI	0,173 n=33	0,288 n=35	0,057 n=33	0,136 n=35	0,171 n=31	0,222 n=34	0,098 n=31	0,197 n=34
KA nein	RC	11.4	17.4	13.5	19.0	9.4	15.0	14.1	25.0
	RI	0,205 n=61	0,313 n=62	0,085 n=61	0,117 n=62	0,176 n=58	0,281 n=61	0,083 n=58	0,153 n=61
PDA ja	RC	10.8	17.5	12.9	20.9	10.2	13.9	15.4	28.3
	RI	0,188 n=43	0,306 n=42	0,078 n=43	0,126 n=42	0,170 n=40	0,234 n=42	0,092 n=40	0,169 n=42
PDA nein	RC	10.7	16.2	11.6	19.7	8.7	13.6	15.0	28.5
	RI	0,202 n=51	0,303 n=55	0,074 n=51	0,125 n=55	0,177 n=49	0,273 n=55	0,090 n=49	0,171 n=55

RC = repeatability coefficient, RI = repeatability index, LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, ; LPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der LPEP; LPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der LPEP; entsprechend für LVET; RPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der RPEP; RPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der RPEP; entsprechend für RVET, Maximum Rot, Minimum Grün, KA = Katecholamine; PDA = persistierender Ductus arteriosus

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2

Im Rahmen der Analyse der reproduzierbaren STI-Messungen sind keine Minimal/Maximalwerte in den Subgruppen der Kategorien SST, PDA, Geschlecht, Sauerstofftherapie oder Beatmungstherapie angesiedelt (siehe Tabelle 10-13).

Es gilt zu sagen, dass auch bei Betrachtung der RI-Werte insgesamt keine bedeutenden Tendenzen zwischen den Subgruppen einer Kategorie feststellbar waren (siehe Tabelle 9-13). Jedoch zeigten insbesondere die interindividuellen LPEP-Messwerte (LPEP1-3) teilweise Schwankungen >10% innerkategorial (fettgedruckt, siehe APGAR Tabelle 10, Seite 35, siehe Geschlecht Tabelle 12, Seite 37). Dies gilt auch bzgl. der RPEP1-3 und der RVET1-3 Messwerte (APGAR Tabelle 10). Hierbei ist diese hohe Schwankung jedoch nur zwischen APGAR <5/ APGAR 5-7 festzustellen.

Tabelle 12 Repeatability coefficient/Repeatability index der Differenzen aller Messwerte gefiltert nach Geschlecht

		PEP Links		VET Links		PEP Rechts		VET Rechts	
		LPEP1-2	LPEP1-3	LVET1-2	LVET1-3	RPEP1-2	RPEP1-3	RVET1-2	RVET1-3
<b>Männlich</b>	<b>RC</b>	10.8	13.8	12.4	21.7	9.2	14.2	14.6	29.7
	<b>RI</b>	0.199	<b>0,254</b>	0,078	0,136	0,170	0,263	0,089	0,180
		n=56	n=57	n=56	n=57	n=53	n=55	n=53	n=55
<b>Weiblich</b>	<b>RC</b>	10.6	20.3	11.1	18.1	9.9	14.2	15.2	26.7
	<b>RI</b>	0,189	<b>0,382</b>	0,068	0,112	0,181	0,261	0,094	0,158
		n=38	n=40	n=38	n=40	n=36	n=40	n=36	n=40

RC = repeatability coefficient, RI = repeatability index, LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, ; LPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der LPEP; LPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der LPEP; entsprechend für LVET; RPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der RPEP; RPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der RPEP; entsprechend für RVET, Maximum Rot, Minimum Grün

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2

Tabelle 13 Repeatability coefficient/Repeatability index der Differenzen aller Messwerte gefiltert nach O<sub>2</sub>-Gabe/Beatmung

		PEP Links		VET Links		PEP Rechts		VET Rechts	
		LPEP1-2	LPEP1-3	LVET1-2	LVET1-3	RPEP1-2	RPEP1-3	RVET1-2	RVET1-3
<b>O2 ja</b>	<b>RC</b>	10.9	16.2	11.7	17.4	9.4	13.1	14.9	27.2
	<b>RI</b>	0,202	0,298	0,073	0,109	0,177	0,241	0,090	0,165
		n=72	n=75	n=72	n=75	n=70	n=74	n=70	n=74
<b>O2 nein</b>	<b>RC</b>	9.7	18.8	12.8	27.6	9.6	17.7	16.2	32.0
	<b>RI</b>	0,169	0,326	0,077	0,167	0,169	0,311	0,093	0,186
		n=22	n=22	n=22	n=22	n=19	n=21	n=19	n=21
<b>Beatmung ja</b>	<b>RC</b>	10.7	18.2	12.0	21.3	9.5	12.8	15.4	31.9
	<b>RI</b>	0,204	0,344	0,075	0,130	0,182	0,246	0,092	0,190
		n=42	n=43	n=42	n=43	n=38	n=41	n=38	n=41
<b>Beatmung nein</b>	<b>RC</b>	10.7	15.6	12.4	18.9	9.5	14.7	15.1	25.4
	<b>RI</b>	0,188	0,275	0,077	0,117	0,171	0,265	0,091	0,154
		n=52	n=53	n=52	n=53	n=51	n=53	n=51	n=53

RC = repeatability coefficient, RI = repeatability index, LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, ; LPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der LPEP; LPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der LPEP; entsprechend für LVET; RPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der RPEP; RPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der RPEP; entsprechend für RVET, Maximum Rot, Minimum Grün, O<sub>2</sub> = Sauerstoff

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2

## 5.5 Inter-/Intraindividuelle Wiederholbarkeit anhand des RC und dem RI bei gleichbleibenden Voraussetzungen hinsichtlich HF und Agitiertheit

### Linksventrikuläre Messwerte (siehe Tabelle 14, Seite 40)

Der intraindividuelle RC für die LPEP lag bei 11,0 der RI bei 0,193 (erste Zeile). Die Schwankungsbreite der Messwerte lag folglich in einem Bereich von bis zu 19,3% normiert auf den durchschnittlichen Messwert. Bei interindividueller Messung lagen diese Werte in einem höheren Bereich (RC=15,4; RI=0,275), es gab Schwankungen um 28%.

Für die LVET-Messungen lag der RC für intraindividuelle Erhebungen bei 11,8, der RI lag deutlich niedriger bei 0,074. Die Schwankungsbreite von 7,4% ist damit erneut akzeptabel. Für interindividuelle Messungen lagen diese Werte jetzt nur minimal höher (RC=13; RI=0,08).

### Rechtsventrikuläre Messwerte (siehe Tabelle 14, Seite 40)

Der intraindividuelle RC für die RPEP lag bei 10,1, der RI bei 0,178 (erste Zeile). Die Schwankungsbreite der Messwerte lag damit in einem Bereich von bis zu 17,8% relativ zum Mittelwert der Messwerte. Bei interindividueller Messung lagen diese Werte abermals höher (RC=12,7; RI=0,229) mit Schwankungen um 23%.

Die RI-Werte für die RVET waren niedriger, hier lag der RC für intraindividuelle Messungen bei 16,9, der RI bei 0,102, was einem akzeptablen Umfang der Abweichungen von bis zu 10,2% entspricht. Für die intraindividuellen Messungen lagen diese Werte erneut höher (RC=31,7; RI=0,191).

Im Vergleich zu den intraindividuellen Mehrfachmessungen waren die interindividuell ermittelten Messwerte weniger reproduzierbar, obwohl sich in dieser Untersuchung die interindividuellen LVET-Messungen kaum von den intraindividuellen unterschieden. Die Ergebnisse zeigen, dass unter HF-Differenzen von maximal 5% und einer VAS-Abweichung von maximal 10% nicht nur die intraindividuellen Werte der LVET (LVET1-2) und RVET (RVET1-2) gut reproduzierbar sind, sondern auch die interindividuelle Erhebung der LVET (LVET1-3).

Hinsichtlich der PEP-Messungen treten einmal mehr höhere Schwankungsbreiten auf. Bei den linksventrikulären interindividuellen Mehrfachmessungen sogar bis zu 27,5% (LPEP1-3), was die Zweifel an der generellen Reproduzierbarkeit der PEP-Messungen verstärkt.

Erneut untersuchten wir die Reproduzierbarkeit abhängig von bestimmten Einflussfaktoren nach dem in 4.4.3 bereits beschriebenen Schema.

#### Erfahrungslevel (siehe Tabelle 14, Seite 40)

Der RI für die LPEP (Konstellation EL2/1 und EL2/2) lag wie in der ersten Untersuchung in einem Bereich von ca. 23-25%. Die RI-Werte für die RPEP (ca.18-23%) und die LVET (ca 7-10%) waren im Vergleich zur ersten Untersuchung etwas niedriger, die Werte der RVET-Messungen (ca. 17-21%) etwas höher.

Die Untersuchungen durch zwei Untersucher mit dem EL 1 (n=2/n=4), sowie die Untersuchungen mit der Konstellation EL1/2 (n=6) sind statistisch nicht verwertbar.

Die Erfahrung der einzelnen Untersucher zeigt erneut keinen eindrücklichen Einfluss auf die Wiederholbarkeit der STI-Messungen. Die besten RI-Werte wurden damit abermals für die intraindividuellen LVET-Messungen erreicht.

Tabelle 14 Repeatability coefficient/Repeatability index der Differenzen der Messwerte und gefiltert nach Erfahrungslevel, HF und VAS gleichbleibend

		PEP Links		VET Links		PEP Rechts		VET Rechts	
		LPEP1-2	LPEP1-3	LVET1-2	LVET1-3	RPEP1-2	RPEP1-3	RVET1-2	RVET1-3
Alle	RC	11	15.4	11.8	13	10.1	12.7	16.9	31.7
	RI	<b>0.193</b> n=43	<b>0.275</b> n=44	<b>0,074</b> n=43	<b>0,08</b> n=44	<b>0,178</b> n=40	<b>0,229</b> n=42	<b>0.102</b> n=40	<b>0,191</b> n=42
EL 1/1	RC	/	12.1	/	3,6	/	14.1	/	84.8
	RI		0,249 n=4		0,023 n=4		0,259 n=2		0,463 n=2
EL 1/2	RC	/	27.1	/	14.5	/	20.3	/	23.5
	RI		0.424 n=6		0,084 n=6		0,351 n=6		0,135 n=6
EL 2/1	RC	/	12.7	/	9.8	/	9.6	/	33.8
	RI		0,232 n=15		<b>0,065</b> n=15		0,182 n=15		0,212 n=15
EL 2/2	RC	/	13.8	/	16	/	12.9	/	27.6
	RI		0,246 n=19		0,099 n=19		0,226 n=19		0,165 n=19

RC = repeatability coefficient, RI = repeatability index, LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, ; LPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der LPEP; LPEP1-3 = Messung 1 Untersucher 1 – Messung 3 Untersucher 2 der LPEP; entsprechend für LVET; RPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der RPEP; RPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der RPEP; entsprechend für RVET, Maximum Rot, Minimum Grün, EL = Erfahrungslevel; EL 1/1 = beide Untersucher hatten Erfahrungslevel 1, EL 1/2= erster Untersucher hatte EL 1, zweiter Untersucher hatte EL 2 usw.

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2

Untersuchung der ausreichend reproduzierbaren intraindividuell erhobenen VETs (LVET1-2; RVET1-2), sowie der interindividuell erhobenen LVET (LVET1-3)

Geschlecht (siehe Tabelle 15, Seite 41)

Die RI-Werte für die intraindividuellen Messwertwiederholungen lagen sowohl für LVET (6,8% weiblich vs. 8,5% männlich), als auch für RVET (9,4% vs. 10,7%, respektive) für die Mädchen geringfügig niedriger. Alle Werte waren akzeptabel. Die interindividuellen LVET-Messergebnisse waren für männliche Probanden (9,4%) etwas niedriger als für weibliche Probanden (11,2%), auch hier war der Unterschied marginal und beide Werte waren grenzwertig akzeptabel.

Tabelle 15 Repeatability coefficient/Repeatability index der Differenzen der Messwerte gefiltert nach Geschlecht, HF und VAS gleichbleibend

		PEP Links		VET Links		PEP Rechts		VET Rechts	
		LPEP1-2	LPEP1-3	LVET1-2	LVET1-3	RPEP1-2	RPEP1-3	RVET1-2	RVET1-3
Männlich	RC	12	13.8	13.1	14.45	9.2	9.9	17.2	35.4
	RI	0,214	0,248	0,085	0,094	0,161	0,176	0,107	0,220
		n=25	n=25	n=25	n=25	n=25	n=55	n=53	n=55
Weiblich	RC	9.8	17.7	11.05	18.08	9.90	14.16	15.77	26.74
	RI	0,169	0,313	0,068	0,112	0,181	0,261	0,094	0,158
		n=18	n=19	n=18	n=19	n=18	n=19	n=18	n=19

RC = repeatability coefficient, RI = repeatability index, LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, ; LPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der LPEP; LPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der LPEP; entsprechend für LVET; RPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der RPEP; RPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der RPEP; entsprechend für RVET, Maximum Rot, Minimum Grün

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2

### Gestationsalter (Schwangerschaftstage-SST, siehe Tabelle 16, Seite 43)

Bei intraindividuellem Messwertwiederholung der LVET zeigt sich für Kinder unter 28 SSW (<196 SST) ein akzeptabler RI von 9,8%. Dieser Wert war für reifere Kinder (>28 SSW) geringfügig niedriger (8,1%), wobei beide Werte im Toleranzbereich lagen.

Ein grenzwertiger RI liegt bei Betrachtung der LVET1-2 in der Subgruppe <196 SST (<28.SSW), (RI=9,8%, rot markiert). Für <259 SST (FG) zeigt sich ein akzeptabler Wert von RI = 8,1%.

Bezüglich der LVET1-3 (RI-SST<259 SST=8,4%; RI-SST<196 SST=8,8%) und der RVET1-2 (RI-SST<259 SST=10,9%; RI-SST<196 SST=10,7%) können die Werte der Subgruppen als gleich angenommen werden.

### Geburtsgewicht (siehe Tabelle 16, Seite 43)

Unter den akzeptablen Werten der LVET1-2, zeigten die Untersuchungen der Kinder mit einem GG < 1500 g einen grenzwertigen RI (RI=9,5%, rot markiert).

Wiesen die NG ein GG ≥1500g auf, war der RI (RI=4,5%, grün markiert) äußerst zufriedenstellend und nochmal ein Prozent höher als bei der ersten ungefilterten Betrachtung der Gesamtkohorte (Vergleich Abschnitt 5.4).

Die Untersuchung der RVET1-2-Messungen hinsichtlich eines möglichen Einflusses des GG kommt zu selbigem Ergebnis, der RI der Subgruppe  $GG < 1500g$  liegt sogar knapp außerhalb des festgelegten Normbereiches (RI=12,7%, rot markiert). Der RI für  $GG \geq 1500g$  ist durchaus akzeptabel (RI=7,2%, grün markiert). Bezüglich LVET1-3 gleichen sich die Werte zunehmend an (RI- $GG < 1500g$  = 9,1%, RI-  $GG \geq 1500g$  = 7,3%).

#### APGAR (siehe Tabelle 16, Seite 43)

Bei Betrachtung der RVET1-2 fällt weiterhin auf, dass die Werte der Messungen bei Kindern mit einem 5-Minuten-APGAR von 5-7 (RI=11,9%) zu den schlechtesten/grenzwertig akzeptablen RIs gehören (rot markiert). Die Untersuchungen der Kinder mit einem  $APGAR \geq 8$  (RI=6,7%) zeigten die besten Werte (grün markiert).

Bezüglich LVET1-2 (RI-APGAR 5-7=8,8%, RI-APGAR  $\geq 8$ =6,1%) und LVET1/3 (RI-APGAR 5-7=7,1%, RI-APGAR  $\geq 8$ =7,5%) gleichen sich die Werte an.

Die Werte der Subgruppe mit einem  $APGAR < 5$  ist bei zu niedriger Kohortenzahl (n=10) statistisch nicht verwertbar.



Tabelle 16 Repeatability coefficient/Repeatability index der Differenzen der Messwerte gefiltert nach SST/GG/APGAR, HF und VAS gleichbleibend

		PEP Links		VET Links		PEP Rechts		VET Rechts	
		LPEP1-2	LPEP1-3	LVET1-2	LVET1-3	RPEP1-2	RPEP1-3	RVET1-2	RVET1-3
SST<259 (37. SSW)	RC	10.9	12.2	12.5	13	10.2	11.3	17.8	32.2
	RI	0,197	0,226	0,081	0,084	0,194	0,216	0,109	0,198
		n=34	n=35	n=34	n=34	n=31	n=33	n=31	n=33
SST<196 (28. SSW)	RC	13.2	13.8	14.7	13	10	14.2	17	14.2
	RI	0,251	0,266	0,098	0,088	0,210	0,296	0,107	0,091
		n=16	n=16	n=16	n=16	n=13	n=14	n=13	n=14
GG<1500 (g)	RC	11.1	12.5	14.2	13.6	9.8	12.1	20.2	37
	RI	0,203	0,234	0,095	0,091	0,198	0,245	0,127	0,234
		n=24	n=25	n=24	n=25	n=21	n=23	n=21	n=23
GG≥1500 (g)	RC	10.2	17.8	7.6	12.5	10.5	13.6	12.6	20.6
	RI	0,172	0,299	0,045	0,073	0,162	0,216	0,072	0,117
		n=19	n=19	n=19	n=19	n=19	n=19	n=19	n=19
APGAR <5	RC	14.1	17.6	4	23.8	12.7	15.9	8.5	36.9
	RI	0,253	0,330	0,027	0,163	0,226	0,296	0,056	0,248
		n=4	n=4	n=4	n=4	n=2	n=3	n=2	n=3
APGAR 5-7	RC	11.6	12.2	13.6	11	10	11.8	19.6	24.8
	RI	0,214	0,226	0,088	0,071	0,189	0,226	0,119	0,151
		n=24	n=24	n=24	n=24	n=23	n=23	n=23	n=23
APGAR ≥8	RC	8.8	17.6	10.1	12.4	10.4	14	11.4	33.4
	RI	0,143	0,294	0,061	0,075	0,166	0,230	0,067	0,193
		n=15	n=16	n=15	n=16	n=15	n=16	n=15	n=16

RC = repeatability coefficient, RI = repeatability index, LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, ; LPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der LPEP; LPEP1-3 = Messung 1 Untersucher 1 – Messung 3 Untersucher 2 der LPEP; entsprechend für LVET; RPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der RPEP; RPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der RPEP; entsprechend für RVET, Maximum Rot, Minimum Grün, SST = Schwangerschaftstage; GG = Geburtsgewicht

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2

### Katecholamine (siehe Tabelle 17, Seite 44)

Die RIs der interindividuell wiederholt gemessenen LVET1-3 zeigten bei den Kindern, die KAS erhielten, eine grenzwertig akzeptable Wiederholbarkeit (RI=9,4%, rot markiert). Der beste RI-Wert der LVET1-3 bezieht sich auf die Subgruppe der Kinder ohne Katecholamingabe (RI=6,7%, grün markiert).

Bezüglich LVET1-2 (RI-KA ja=6,1%, RI-KA nein=8,5%) und RVET1/2 (RI-KA ja=10,7%, RI-KA nein = 9,8%) gleichen sich die Werte an.

## PDA (siehe Tabelle 17)

Der RI der LVET1-3 zeigte bei den Kindern, die zum Zeitpunkt der Untersuchung einen PDA vorwiesen, eine grenzwertig akzeptable Wiederholbarkeit (RI=9,6%, rot markiert). Der RI-Wert ohne PDA ist akzeptabel (RI=7,3%).

Bezüglich LVET1-2 (RI-PDA ja=7,2%, RI-PDA nein=7,8%) und RVET1-2 (RI-PDA ja=9%, RI-PDA nein = 11%) gleichen sich die Werte an.

*Tabelle 17 Repeatability coefficient/Repeatability index der Differenzen der Messwerte gefiltert nach KA-Gabe/PDA, HF und VAS gleichbleibend*

		PEP Links		VET Links		PEP Rechts		VET Rechts	
		LPEP1-2	LPEP1-3	LVET1-2	LVET1-3	RPEP1-2	RPEP1-3	RVET1-2	RVET1-3
KA ja	RC	10.6	15.6	9.9	15.3	10.5	12	17.8	33.8
	RI	0,193	0,286	0,061	<b>0,094</b>	0,181	0,212	0,107	0,2
		n=21	n=22	n=21	n=22	n=19	n=21	n=19	n=21
KA nein	RC	11.6	15.4	13.4	10.6	9.8	15.02	16.4	30
	RI	0,197	0,266	0,085	<b>0,067</b>	0,176	0,252	0,098	0,180
		n=22	n=22	n=22	n=22	n=21	n=21	n=21	n=21
PDA ja	RC	12.2	20.9	12	16	10	11.8	15	30.7
	RI	0,202	<b>0,353</b>	0,072	<b>0,096</b>	0,152	0,186	0,090	0,185
		n=17	n=17	n=17	n=17	n=16	n=17	n=16	n=17
PDA nein	RC	10.5	10.8	11.8	11.2	10.4	12.7	18.2	32.8
	RI	0,192	<b>0,200</b>	0,078	0,073	0,206	0,253	0,110	0,197
		n=26	n=27	n=26	n=27	n=24	n=25	n=24	n=25

RC = repeatability coefficient, RI = repeatability index, LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, ; LPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der LPEP; LPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der LPEP; entsprechend für LVET; RPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der RPEP; RPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der RPEP; entsprechend für RVET, Maximum Rot, Minimum Grün, KA = Katecholamine; PDA = persistierender Ductus arteriosus

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2

Hinsichtlich der Kategorien Sauerstofftherapie und Beatmungstherapie sind darüber hinaus im Rahmen der Analyse der reproduzierbaren STI-Messungen keine Maxima/Minima erkennbar (siehe Tabelle 18, Seite 45).

Bei der Analyse sämtlicher RI Werte wurden keine klaren Tendenzen zwischen den Subgruppen detektiert. Es traten teilweise deutliche Abweichungen der RIs auf. Innerkategoriale Schwankungen zwischen den Subgruppen >10% betrafen LPEP1-3 (siehe PDA ja/nein Tabelle 17, O2-Therapie Tabelle 18) und RPEP1-2 (siehe O2-Therapie Tabelle 18). Die RVET1-3 zeigt erhebliche Schwankungen in der Kategorie GG (siehe Tabelle 16).

Tabelle 18 Repeatability coefficient/Repeatability index der Differenzen der Messwerte gefiltert nach O<sub>2</sub>-Gabe/Beatmung, HF und VAS gleichbleibend

		PEP Links		VET Links		PEP Rechts		VET Rechts	
		LPEP1-2	LPEP1-3	LVET1-2	LVET1-3	RPEP1-2	RPEP1-3	RVET1-2	RVET1-3
O <sub>2</sub> ja	RC	10.6	12.4	10.8	12.8	10.6	11	17.5	30.5
	RI	0,190 n=37	<b>0,223</b> n=37	0,068 n=37	0,081 n=37	<b>0,191</b> n=35	<b>0,201</b> n=36	0,106 n=35	0,166 n=36
O <sub>2</sub> nein	RC	13.8	27.5	13.2	14.7	5.6	22.2	12.8	41.2
	RI	0,213 n=6	<b>0,462</b> n=7	0,080 n=6	0,088 n=7	<b>0,086</b> n=5	<b>0,369</b> n=6	0,073 n=5	0,232 n=6
Beatmung ja	RC	12	14	12.7	13.6	10.1	14	16.6	31.2
	RI	0,220 n=25	0,255 n=25	0,080 n=25	0,085 n=25	0,183 n=22	0,254 n=23	0,099 n=22	0,186 n=23
Beatmung nein	RC	9.8	17	10.6	12.6	10.3	11.2	17.8	32.7
	RI	0,164 n=18	0,293 n=19	0,066 n=18	0,079 n=19	0,176 n=18	0,199 n=19	0,108 n=18	0,196 n=19

RC = repeatability coefficient, RI = repeatability index, LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, ; LPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der LPEP; LPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der LPEP; entsprechend für LVET; RPEP1-2 = Messung 1 Untersucher 1-Messung 2 Untersucher 1 der RPEP; RPEP1-3 = Messung 1 Untersucher 1 - Messung 3 Untersucher 2 der RPEP; entsprechend für RVET, Maximum Rot, Minimum Grün, O<sub>2</sub>= Sauerstoff

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2

## 5.6 Inter-/Intraindividuelle Wiederholbarkeit anhand des Cohens Kappa Koeffizienten

Für diese Analyse haben wir nur die Frühgeborenen ausgewählt, da ca. 87% unserer Kohorte Frühgeborene waren und für Reifgeborene andere Normbereiche gelten.

Im Hinblick auf die gesamte Kohorte ist die Reliabilität erneut regelhaft bei den intraindividuellen Mehrfachmessungen von STIs größer als die der interindividuell wiederholten Messungen.

Die qualitative Beurteilung der Werte mittels Cohens Kappa Koeffizienten bestätigte zunächst die ausreichende Reproduzierbarkeit für die LVET1/2 und RVET1/2, ermittelt in den RC/RI-Untersuchungen. Es zeigte sich eine definitionsgemäß beachtliche Reliabilität (siehe Tabelle 1, Seite 21) hinsichtlich der intraindividuellen Übereinstimmung (LVET1/2=0,621 und RVET1/2=0,724, Tabelle 19, Seite 47, erste Zeile).

Bei Analyse der interindividuellen Reliabilität liegen diese Werte aber schon erheblich niedriger (LVET1/3=0,37, ausreichende Reliabilität; RVET1/3=0,486, moderate Reliabilität).

Erwartungsgemäß waren auch bei der Cohens-Kappa-Analyse die Werte für die PEP in der Regel niedriger als die der VET. In allen Fällen zeigte sich eine moderate Reliabilität (LPEP1/2=0,509; LPEP1/3=0,445; RPEP1/2=0,568; RPEP1/3=0,413). Der Kappa-Wert für LPEP1/3 war ausnahmsweise höher als für LVET1/3.

### Erfahrungslevel (siehe Tabelle 19, Seite 47)

Bei den linksventrikulären PEP-Messungen war eine Reliabilität im ausreichenden bis moderaten Bereich festzustellen (LPEP1/3:  $EL2/2=0,259 < EL2/1=0,468 < EL1/2=0,477$ ), die LVET-Messungen erbrachten ebenfalls Werte mit ausreichender bis moderater Übereinstimmung (LVET1/3:  $EL2/1=0,325 < EL2/2=0,342 < EL1/2=0,426$ ).

Auch rechtsventrikulär waren die Messergebnisse für die RPEP im ausreichenden bis moderaten Bereich angesiedelt, bezüglich der Verlässlichkeit der Untersuchungen (RPEP1/3:  $EL1/2=0,328 < EL2/2=0,385 < EL2/1=0,445$ ). Die größten Schwankungen traten bei den RVET-Messungen auf, von ausreichender bis beachtlicher Übereinstimmung (RVET1/3:  $EL2/1=0,225 < EL1/2=0,595 < EL2/2=0,633$ ).

Unter der Gegebenheit, dass zwei Untersucher EL = 2 die Daten erhoben haben, ist folglich, bis auf die RVET-Messungen, keine höhere Reproduzierbarkeit erreicht worden. Die Kappa Werte waren teilweise sogar niedriger (LPEP1/3).

EL1/1 wurde bei zu niedriger Anzahl an Messungen und dadurch teilweise nicht durchführbarer Analyse nicht beachtet (n=2/n=4).

Die Analyse der erhobenen STIs, selektiert nach der Voraussetzung, dass bei beiden Messungen die EL der Untersucher gleich waren, erbrachte zusammenfassend keinen entscheidenden Unterschied gegenüber den Gruppen mit Beobachtern unterschiedlicher Expertise. In sämtlichen Subgruppen ist eine ausreichende bis maximal beachtliche Übereinstimmung erzielt worden.

*Tabelle 19 Cohens Kappa Koeffizient gesamte Kohorte und gefiltert nach Expertise*

	PEP Links		VET Links		PEP Rechts		VET Rechts	
Cohens Kappa	LPEP1/2	LPEP1/3	LVET1/2	LVET1/3	RPEP1/2	RPEP1/3	RVET1/2	RVET1/3
<b>Alle</b>	<b>0,509</b> n=83	<b>0,445</b> n=86	<b>0,621</b> n=83	<b>0,370</b> n=85	<b>0,568</b> n=78	<b>0,413</b> n=84	<b>0,724</b> n=78	<b>0,486</b> n=84
<b>EL 1/1</b>	/	1 n=4	/	1 n=4	/	1 n=2	/	n.m. n=2
<b>EL 1/2</b>	/	0,477 n=15	/	0,426 n=14	/	0,328 n=15	/	0,595 n=15
<b>EL 2/1</b>	/	0,468 n=33	/	0,325 n=33	/	0,445 n=33	/	0,225 n=33
<b>EL 2/2</b>	/	0,259 n=34	/	0,342 n=34	/	0,385 n=34	/	0,633 n=34

*LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, LPEP1/2 = Übereinstimmung Betrachtung LPEP Untersucher 1, 1. vs. 2. Messung; LPEP1/3 = Übereinstimmung Betrachtung LPEP Untersucher 1 vs. Untersucher 2; LVET1/2 = Übereinstimmung Betrachtung LVET Untersucher 1, 1. vs. 2. Messung; LVET1/3 = Übereinstimmung Betrachtung LVET Untersucher 1 vs. Untersucher 2; entsprechend für rechtsventrikuläre Werte; EL = Erfahrungslevel; EL 1/1 = beide Untersucher hatten Erfahrungslevel 1, EL 1/2= erster Untersucher hatte EL 1, zweiter Untersucher hatte EL 2 usw.; n.m. statistische Analyse nicht möglich*

*Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2*

### Geschlecht (Tabelle 20, Seite 48)

Die intraindividuellen Messergebnisse, bis auf die LPEP-Messungen bei weiblichen Individuen (LPEP1/2: weiblich=0,513 < LPEP1/3: weiblich=0,67) waren durchgehend höher als die interindividuellen.

Dabei war die Übereinstimmung der LVET-Messungen und auch der RVET-Messungen regelhaft höher als die der PEP-Messungen (LPEP1/2: männlich=0,5; weiblich=0,513 < LVET1/2: männlich=0,621; weiblich=0,591; RPEP1/2: männlich=0,556; weiblich=0,59; < RVET1/2: männlich=0,739; weiblich=0,695).

Auffällig hoch ist die Abweichung der Kappa-Werte der LPEP1/3 in der Subgruppenanalyse (LPEP1/3: männlich=0,293; weiblich=0,67). Sonst gleichen sich die Werte der Subgruppen weitestgehend an.

Die rechtsventrikulären STI-Messungen zeigten in der Kategorie Geschlecht größtenteils eine etwas höhere Verlässlichkeit (bis auf LPEP1/3: weiblich=0,67 > RPEP1/3: weiblich=0,319).

*Tabelle 20 Cohens Kappa Koeffizient gefiltert nach Geschlecht*

	PEP Links		VET Links		PEP Rechts		VET Rechts	
Cohens Kappa	LPEP1/2	LPEP1/3	LVET1/2	LVET1/3	RPEP1/2	RPEP1/3	RVET1/2	RVET1/3
<b>Männlich</b>	0,500 n=52	<b>0,293</b> n=53	0,621 n=52	0,364 n=53	0,556 n=49	0,469 n=51	0,739 n=49	0,507 n=51
<b>Weiblich</b>	0,513 n=31	<b>0,670</b> n=33	0,591 n=31	0,346 n=33	0,590 n=29	0,319 n=33	0,695 n=29	0,455 n=33

*LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, LPEP1/2 = Übereinstimmung Betrachtung LPEP Untersucher 1, 1. vs. 2. Messung; LPEP1/3 = Übereinstimmung Betrachtung LPEP Untersucher 1 vs. Untersucher 2; LVET1/2 = Übereinstimmung Betrachtung LVET Untersucher 1, 1. vs. 2. Messung; LVET1/3 = Übereinstimmung Betrachtung LVET Untersucher 1 vs. Untersucher 2; entsprechend für rechtsventrikuläre Werte*

*Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2*

### Gestationsalter (Schwangerschaftstage-SST, siehe Tabelle 21, Seite 51)

Hinsichtlich der Schwangerschaftsdauer waren die Werte der intraindividuellen Messungen nicht durchgehend höher als die interindividuellen. Die LPEP1/3 zeigte in beiden Subgruppen (SST<224=0,517; SST<196=0,535) eine größere Reliabilität als die LPEP1/2 (SST<224=0,471; SST<196=0,465).

Die Übereinstimmung der intraindividuellen LVET-Messungen und auch der intraindividuellen RVET-Messungen war erneut durchgehend höher als die der PEP-Messungen (LPEP1/2: SST<224=0,471; SST<196=0,465 < LVET1/2: SST<224=0,612; SST<196=0,613; RPEP1/2: SST<224=0,587; SST<196=0,571 < RVET1/2: SST<224=0,63; SST<196=0,586).

Bei den interindividuellen STI-Messungen gilt dies nur hinsichtlich der rv Erhebungen (RPEP1/3: SST<224=0,343; SST<196=0,342 < RVET1/3: SST<224=0,42; SST<196=0,481).

Die Kappa-Werte der jeweiligen STI-Messungen glichen sich zwischen den Subgruppen an.

#### Geburtsgewicht (siehe Tabelle 21, Seite 51)

Hinsichtlich des Geburtsgewichtes waren die Werte der intraindividuellen Messungen nicht durchgehend größer als die interindividuellen. Die jeweils höheren Kappa-Werte der LPEP1/3 (GG<1500g=0,536) und der RPEP1/3 (GG≥1500g=0,61) überstiegen die entsprechenden intraindividuellen Werte.

Die Übereinstimmung der intraindividuellen LVET-Messungen und auch der intraindividuellen RVET-Messungen war erneut etwas höher als die der PEP-Messungen (LPEP1/2: GG<1500g=0,493; GG≥1500g=0,548 < LVET1/2: GG<1500g=0,612; GG≥1500g=0,558; RPEP1/2: GG<1500g=0,569; GG≥1500g=0,523 < RVET1/2: GG<1500g=0,606; GG≥1500g=0,904).

Bei den interindividuellen STI-Messungen gilt dies nur hinsichtlich der rv Erhebungen (RPEP1/3: GG<1500g=0,327; GG≥1500g=0,61 < RVET1/3: GG<1500g=0,388; GG≥1500g=0,621).

Starke Abweichungen der Kappa-Werte der Subgruppen zeigten vor allem die LPEP1/3 (GG<1500g=0,536; GG≥1500g=0,184), die RPEP1/3 (GG<1500g=0,327; GG≥1500g=0,61), die RVET1/2 (GG<1500g=0,606; GG≥1500g=0,904), sowie die RVET1/3 (GG<1500g=0,388; GG≥1500g=0,621).

Die Kappa-Werte einiger STI-Messungen glichen sich zwischen den Subgruppen an (LPEP1/2, LVET1/2, LVET1/3, RPEP1/2).

#### APGAR (siehe Tabelle 21, Seite 51)

Hinsichtlich des APGARs waren die Kappa-Werte der intraindividuellen Messungen ca. gleich oder höher als die der interindividuellen.

Die Übereinstimmung der intraindividuellen LVET-Messungen und auch der intraindividuellen RVET-Messungen war erneut höher als die der entsprechenden PEP-Messungen (bis auf APGAR-Wert  $\geq 8$ :  $\underline{RPEP1/2}=0,645 > \underline{RVET1/2}=0,528$ ).

Die Werte der Kinder mit einem 5-Minuten-APGAR von 5-7 waren regelhaft niedriger als die Werte der Kinder mit einem APGAR-Wert  $\geq 8$  (bis auf APGAR 5-7:  $\underline{RVET1/2}=0,825$ ;  $\underline{RVET1/3}=0,681 > \text{APGAR-Wert } \geq 8$ :  $\underline{RVET1/2}=0,528$ ;  $\underline{RVET1/3}=0,253$ ).

Für APGAR 5-7 waren die Kappa-Werte der rv STI-Messungen stets höher als die der lv STI-Messungen.

#### Lebenstag (siehe Tabelle 21, Seite 51)

Für  $LT < 5$  waren die Kappa-Werte der intraindividuellen Messungen durchweg größer als die der interindividuellen. Dies gilt nicht für  $LT \geq 5$ , in der Subgruppe glichen sich die Kappa-Werte der Intra- und interindividuellen Messungen an, für die lv STI-Messungen waren die interindividuellen Messungen sogar diskret höher.

Die Übereinstimmung der LVET-Messungen und auch der RVET-Messungen war erneut höher als die der PEP-Messungen (bis auf  $LT < 5$ :  $\underline{LPEP1/3}=0,586 > \underline{LVET1/3}=0,218$ ;  $\underline{RPEP1/3}=0,475 > \underline{RVET1/3}=0,301$ ).

Sämtliche Kappa-Werte der PEP-Messungen und auch die der intraindividuellen VET-Messungen waren größer für die Kinder, die vor dem 5. LT untersucht wurden. Nur die interindividuellen VET-Messungen waren niedriger ( $< 5.LT$ :  $\underline{LVET1/3}=0,218$ ;  $\underline{RVET1/3}=0,301$ ). Es zeigten sich starke Abweichungen gegenüber der Subgruppe  $\geq 5.LT$  ( $\underline{LVET1/3}=0,531$ ;  $\underline{RVET1/3}=0,632$ ).



Tabelle 21 Cohens Kappa Koeffizient gefiltert nach SST/GG/APGAR und LT

	PEP Links		VET Links		PEP Rechts		VET Rechts	
Cohens Kappa	LPEP1/2	LPEP1/3	LVET1/2	LVET1/3	RPEP1/2	RPEP1/3	RVET1/2	RVET1/3
SST<224	0,471 n=66	0,517 n=69	0,612 n=66	0,314 n=68	0,587 n=61	0,343 n=67	0,630 n=61	0,420 n=67
SST<196	0,465 n=41	0,535 n=42	0,613 n=41	0,375 n=41	0,571 n=37	0,342 n=40	0,586 n=37	0,481 n=39
GG<1500 (g)	0,493 n=64	<b>0,536</b> n=66	0,612 n=64	0,327 n=65	0,569 n=59	<b>0,327</b> n=64	<b>0,606</b> n=59	<b>0,388</b> n=64
GG≥1500 (g)	0,548 n=19	<b>0,184</b> n=20	0,558 n=19	0,316 n=20	0,523 n=19	<b>0,610</b> n=20	<b>0,904</b> n=19	<b>0,621</b> n=20
APGAR 5-7	<b>0,385</b> n=39	<b>0,381</b> n=41	<b>0,546</b> n=39	<b>0,395</b> n=40	<b>0,430</b> n=37	<b>0,400</b> n=40	<b>0,825</b> n=37	<b>0,681</b> n=40
APGAR ≥8	0,559 n=35	0,564 n=36	0,646 n=35	0,452 n=36	0,645 n=34	0,528 n=36	0,528 n=34	0,253 n=36
LT < 5	0,704 n=41	0,586 n=43	0,742 n=41	<b>0,218</b> n=43	0,662 n=36	0,475 n=41	0,748 n=36	<b>0,301</b> n=41
LT ≥ 5	0,237 n=41	0,242 n=42	0,49 n=41	<b>0,531</b> n=42	0,407 n=41	0,341 n=42	0,686 n=41	<b>0,632</b> n=42

LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, LPEP1/2 = Übereinstimmung Betrachtung LPEP Untersucher 1 1. Vs. 2. Messung; LPEP1/3 = Übereinstimmung Betrachtung LPEP Untersucher 1 vs. Untersucher 2; LVET1/2 = Übereinstimmung Betrachtung LVET Untersucher 1 1. Vs. 2. Messung; LVET1/3 = Übereinstimmung Betrachtung LVET Untersucher 1 vs. Untersucher 2; entsprechend für rechtsventrikuläre Werte; SST = Schwangerschaftstage; GG = Geburtsgewicht; LT = Lebenstag;

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2

### Sauerstofftherapie (siehe Tabelle 22, Seite 52)

Die Übereinstimmung der intraindividuellen Messungen war prinzipiell größer (bis auf O<sub>2</sub> nein: LPEP1/2 = 0,564 < LPEP1/3 = 0,717).

Bei Kindern, die zusätzlichen Sauerstoff erhielten, waren VET-Messungen verlässlicher als die PEP-Messungen. Dies gilt nicht für die Kinder ohne Sauerstofftherapie. Die VET-Messungen stimmten durchgehend weniger überein als die PEP-Messungen.

Bis auf die RPEP1/3 (O<sub>2</sub>-Gabe=0,42; keine O<sub>2</sub>-Gabe=0,389) waren sämtliche Kappa-Werte der PEP-Messungen bei Kindern mit O<sub>2</sub>-Insufflation niedriger, die Werte der VET-Messungen lagen durchweg höher.

In den Subgruppenanalysen O<sub>2</sub> ja/nein gab es einige Abweichungen >10%. Besonders auffällig waren die Werte der LPEP1/3, RPEP1/2, RVET1/2 und RVET1/3.

### Beatmung (siehe Tabelle 22)

Bei den beatmeten Kindern zeigten die intraindividuellen STI-Messungen regelhaft verlässlichere Werte.

Hinsichtlich PEP/VET zeigte sich keine Regelhaftigkeit.

Die LPEP1/2, die RPEP1/2 und die RVET1/2 zeigten eine deutlich höhere Übereinstimmung bei den Kindern, die beatmet wurden.

In den Subgruppenanalysen Beatmung ja/nein gab es einige Abweichungen >10%. Besonders auffällig waren die Werte der LPEP1/2, RPEP1/2, RVET1/2 und RVET1/3.

*Tabelle 22 Cohens Kappa Koeffizient gefiltert nach O<sub>2</sub>-Gabe und Beatmung*

	PEP Links		VET Links		PEP Rechts		VET Rechts	
Cohens Kappa	LPEP1/2	LPEP1/3	LVET1/2	LVET1/3	RPEP1/2	RPEP1/3	RVET1/2	RVET1/3
O <sub>2</sub> ja	0,495 n=66	<b>0,382</b> n=69	0,647 n=66	0,399 n=69	<b>0,483</b> n=64	0,420 n=68	<b>0,843</b> n=64	<b>0,599</b> n=68
O <sub>2</sub> nein	0,564 n=17	<b>0,717</b> n=17	0,472 n=17	0,317 n=17	<b>0,890</b> n=14	0,389 n=16	<b>0,255</b> n=14	<b>0,067</b> n=16
Beatmung ja	<b>0,619</b> n=35	0,385 n=36	0,604 n=35	0,312 n=36	<b>0,664</b> n=31	0,366 n=34	<b>0,859</b> n=31	0,559 n=34
Beatmung nein	<b>0,290</b> n=48	0,395 n=49	0,625 n=48	0,461 n=49	<b>0,481</b> n=47	0,483 n=49	<b>0,594</b> n=47	0,433 n=49

LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, LPEP1/2 = Übereinstimmung Betrachtung LPEP Untersucher 1 1. Vs. 2. Messung; LPEP1/3 = Übereinstimmung Betrachtung LPEP Untersucher 1 vs. Untersucher 2; LVET1/2 = Übereinstimmung Betrachtung LVET Untersucher 1 1. Vs. 2. Messung; LVET1/3 = Übereinstimmung Betrachtung LVET Untersucher 1 vs. Untersucher 2; entsprechend für rechtsventrikuläre Werte; O<sub>2</sub> = Sauerstoff

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2

### Katecholamingabe (siehe Tabelle 23, Seite 53)

Bei den Kindern mit Katecholamintherapie zeigten die intraindividuellen STI-Messungen regelhaft verlässlichere Werte.

Die VET-Messungen waren meist verlässlicher als die PEP-Messungen (bis auf KA nein: LPEP1/3=0,591 > LVET1/3=0,357).

Die LPEP1/3, die RPEP1/3 und die RVET1/3 zeigten eine deutlich größere Übereinstimmung bei den Kindern, ohne Katecholamintherapie.

In den Subgruppenanalysen Katecholamingabe ja/nein gab es einige Abweichungen >10%. Besonders auffällig waren die Werte der LPEP1/3, RPEP1/3 und RVET1/3.

#### PDA (siehe Tabelle 23)

Die intraindividuellen VET-Messungen zeigten stets eine größere Übereinstimmung als die interindividuellen. Bis auf die LPEP-Messungen bei Kindern mit PDA (LPEP1/2=0,529; LPEP1/3=0,538) waren auch hier die intraindividuell wiederholt gemessenen STIs verlässlicher.

Die VET-Messungen zeigten regelhaft eine höhere Übereinstimmung als die PEP-Messungen (bis auf PDA ja: LPEP1/3=0,538 > LVET1/3=0,207).

Die VET-Messungen demonstrierten außerdem bis auf die RVET1/3 deutliche Schwankungen bzgl. der Kappa-Werte zwischen den Subgruppen PDA ja/nein.

Die PEP-Messungen gleichen sich hinsichtlich der Verlässlichkeit in den Subgruppen PDA ja/nein weitestgehend an.

*Tabelle 23 Cohens Kappa Koeffizient gefiltert nach KA-Gabe und PDA*

	PEP Links		VET Links		PEP Rechts		VET Rechts	
Cohens Kappa	LPEP1/2	LPEP1/3	LVET1/2	LVET1/3	RPEP1/2	RPEP1/3	RVET1/2	RVET1/3
<b>KA ja</b>	0,592 n=24	<b>0,123</b> n=26	0,746 n=24	0,398 n=26	0,442 n=22	<b>0,117</b> n=25	0,696 n=22	<b>0,274</b> n=25
<b>KA nein</b>	0,445 n=59	<b>0,591</b> n=60	0,562 n=59	0,357 n=59	0,616 n=56	<b>0,540</b> n=59	0,738 n=56	<b>0,587</b> n=59
<b>PDA ja</b>	0,529 n=37	0,538 n=36	<b>0,703</b> n=37	<b>0,207</b> n=35	0,512 n=34	0,401 n=36	<b>0,573</b> n=34	0,500 n=36
<b>PDA nein</b>	0,495 n=46	0,376 n=49	<b>0,549</b> n=46	<b>0,499</b> n=49	0,596 n=44	0,423 n=47	<b>0,804</b> n=44	0,501 n=47

LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, LPEP1/2 = Übereinstimmung Betrachtung LPEP Untersucher 1 1. Vs. 2. Messung; LPEP1/3 = Übereinstimmung Betrachtung LPEP Untersucher 1 vs. Untersucher 2; LVET1/2 = Übereinstimmung Betrachtung LVET Untersucher 1 1. Vs. 2. Messung; LVET1/3 = Übereinstimmung Betrachtung LVET Untersucher 1 vs. Untersucher 2; entsprechend für rechtsventrikuläre Werte; KA = Katecholamine; PDA = persistierender Ductus arteriosus

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2

## 5.7 Inter-/Intraindividuelle Wiederholbarkeit anhand des Cohens Kappa Koeffizienten bei gleichbleibenden Voraussetzungen hinsichtlich HF und Agitiertheit

Bei der Betrachtung der STIs mit annähernd gleichbleibenden Untersuchungsbedingungen hinsichtlich VAS/HF, zeigte sich eine annähernd gleiche oder höhere intraindividuelle Reliabilität im Vergleich zu den Werten der interindividuellen STIs. Nur die LPEP1/2-Werte zeigten eine diskret niedrigere Übereinstimmung als die LPEP1/3-Werte.

Die qualitative Beurteilung der Werte mittels Cohens Kappa Koeffizienten bei gleichbleibender HF/VAS bestätigte abermals die Reproduzierbarkeit aus den RC/RI-Untersuchungen für die LVET1/2 und RVET1/2. Es zeigte sich eine definitionsgemäß moderate bis beachtliche intraindividuelle Reliabilität (siehe Tabelle 1, Seite 21) der VET-Messungen (LVET1/2=0,582 und RVET1/2=0,757; Tabelle 24, Seite 55, erste Zeile).

Die Analyse unter Ausschluss der Untersuchungen mit zu stark abweichender HF / VAS führte bzgl. der RVET1/2 zu einer geringen Zunahme des Kappa-Koeffizienten. Der Wert für die LVET1/2 verringerte sich.

Bei Analyse der interindividuellen Reliabilität liegen die VET-Werte erneut niedriger als die intraindividuellen. Jedoch konnten höhere Kappa-Werte für die interindividuellen Messungen erzielt werden als bei der ersten ungefilterten Analyse (LVET1/3=0,501 und RVET1/3=0,571; Tabelle 24, Seite 55, erste Zeile).

Erwartungsgemäß waren erneut die Werte für die PEP niedriger als die der VET. Die Werte für die PEP-Messungen erniedrigten sich auch insgesamt. Die Kappa-Werte der RPEP1/2 zeigten eine moderate Übereinstimmung, sonst war die Wiederholbarkeit nur ausreichend (LPEP1/2=0,317; LPEP1/3=0,347; RPEP1/2=0,413; RPEP1/3=0,296; Tabelle 24, Seite 55, erste Zeile).

### Erfahrungslevel (siehe Tabelle 24)

Bei den linksventrikulären PEP-Messungen zeigte sich etwas bis ausreichend Übereinstimmung (LPEP1/3:  $EL2/1=0,035$ ;  $EL2/2=0,248$ ). Der Vergleich wiederholter LVET-Messungen war durch eine ausreichende bis moderate Wiederholbarkeit gekennzeichnet (LVET1/3:  $EL2/1=0,371$ ;  $EL2/2=0,598$ ).

Die größte Schwankung trat bei den Messergebnissen für die RPEP auf. Für  $EL2/1$  zeigte sich eine beachtliche Übereinstimmung, während die Werte bei  $EL2/2$  nur etwas übereinstimmten (RPEP1/3:  $EL2/2=0,136 < EL2/1=0,621$ ). Bezüglich der RVET-Messungen ergaben sich ausreichende bis beachtliche Kappa-Werte (RVET1/3:  $EL2/1=0,389 < EL2/2=0,742$ ).

Die Konstellationen  $EL1/2$  und  $EL1/1$  waren aufgrund niedriger Fallzahlen nicht statistisch verwertbar.

Die Analyse der STIs für  $EL2/2$  zeigte bis auf die RPEP1/3 damit eine deutlich höhere Übereinstimmung als die Werte der Konstellation  $EL2/1$ . Die weitere Unterteilung nach Untersucherexpertise deutet in dem Fall tendenziell auf eine Relevanz hinsichtlich der Zuverlässigkeit der Messergebnisse hin. Die Kohortenzahlen sind jedoch niedrig ( $n=11/n=17$ ).

*Tabelle 24 Cohens Kappa Koeffizient bei gleichbleibender HF/VAS und gefiltert nach Expertise*

	PEP Links		VET Links		PEP Rechts		VET Rechts	
Cohens Kappa	LPEP1/2	LPEP1/3	LVET1/2	LVET1/3	RPEP1/2	RPEP1/3	RVET1/2	RVET1/3
Alle	<b>0,317</b> n=34	<b>0,347</b> n=35	<b>0,582</b> n=34	<b>0,501</b> n=35	<b>0,413</b> n=31	<b>0,296</b> n=33	<b>0,757</b> n=31	<b>0,571</b> n=33
EL 1/1	/	n.m. n=3	/	1 n=3	/	n.m. n=1	/	n.m. n=1
EL 1/2	/	0,556 n=4	/	0,2 n=4	/	-0,333 n=4	/	0,5 n=4
EL 2/1	/	0,035 n=11	/	0,371 n=11	/	0,621 n=11	/	0,389 n=11
EL 2/2	/	0,248 n=17	/	0,598 n=17	/	0,136 n=17	/	0,742 n=17

*LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, LPEP1/2 = Übereinstimmung Betrachtung LPEP Untersucher 1 1. Vs. 2. Messung; LPEP1/3 = Übereinstimmung Betrachtung LPEP Untersucher 1 vs. Untersucher 2; LVET1/2 = Übereinstimmung Betrachtung LVET Untersucher 1 1. Vs. 2. Messung; LVET1/3 = Übereinstimmung Betrachtung LVET Untersucher 1 vs. Untersucher 2; entsprechend für rechtsventrikuläre Werte; EL = Erfahrungslevel; EL 1/1 = beide Untersucher hatten Erfahrungslevel 1, EL 1/2 = erster Untersucher hatte EL 1, zweiter Untersucher hatte EL 2 usw., n.m. = statistische Analyse nicht möglich*

*Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2*

### Geschlecht (siehe Tabelle 25)

Hinsichtlich des Geschlechts zeigte sich keine klare Tendenz zwischen den intraindividuellen und den interindividuellen Messungen.

Auffällig ist die Abweichung der Kappa-Werte der PEP-Messungen in der Subgruppenanalyse (LPEP1/2: männlich=0,097; weiblich=0,658, LPEP1/3: männlich=0,152; weiblich=0,592, siehe Tabelle 25). Aber auch die RPEP1/3-, die LVET1/3- und die RVET1/2-Werte zeigten starke Abweichungen zwischen den Subgruppen.

*Tabelle 25 Cohens Kappa Koeffizient bei gleichbleibender HF/VAS gefiltert nach Geschlecht*

	PEP Links		VET Links		PEP Rechts		VET Rechts	
Cohens Kappa	LPEP1/2	LPEP1/3	LVET1/2	LVET1/3	RPEP1/2	RPEP1/3	RVET1/2	RVET1/3
<b>Männlich</b>	<b>0,097</b> n=21	<b>0,152</b> n=21	0,547 n=21	<b>0,396</b> n=21	0,357 n=18	<b>0,395</b> n=19	<b>0,609</b> n=18	0,565 n=19
<b>Weiblich</b>	<b>0,658</b> n=13	<b>0,592</b> n=14	0,610 n=13	<b>0,682</b> n=14	0,475 n=13	<b>0,125</b> n=14	<b>0,843</b> n=13	0,553 n=14

*LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, LPEP1/2 = Übereinstimmung Betrachtung LPEP Untersucher 1 1. Vs. 2. Messung; LPEP1/3 = Übereinstimmung Betrachtung LPEP Untersucher 1 vs. Untersucher 2; LVET1/2 = Übereinstimmung Betrachtung LVET Untersucher 1 1. Vs. 2. Messung; LVET1/3 = Übereinstimmung Betrachtung LVET Untersucher 1 vs. Untersucher 2; entsprechend für rechtsventrikuläre Werte*

*Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2*

### Gestationsalter (Schwangerschaftstage-SST siehe Tabelle 26, Seite 58)

Hinsichtlich der Schwangerschaftsdauer waren die intraindividuellen Messungen nur bezüglich der rv STIs durchgehend höher als die interindividuellen. Die lv STI-Messungen zeigten meist interindividuell eine größere Übereinstimmung (LPEP1/3: SST<224=0,339; SST<196=0,375 > LPEP1/2: SST<224=0,126; SST<196=0,279, LVET1/3: SST<196=0,631 > LVET1/2: SST<196=0,587).

Die Übereinstimmung der VET-Messungen war durchgehend höher als die der PEP-Messungen.

#### Geburtsgewicht (siehe Tabelle 26, Seite 58)

Bis auf LPEP1/2:  $GG < 1500g = 0,158 < \underline{LPEP1/3}$ :  $GG < 1500g = 0,379$  waren die intraindividuellen Messwerte durchgehend gleich oder zeigten eine größere Übereinstimmung.

Die Kappa-Werte der VET-Messungen waren durchgehend höher als die der PEP-Messungen.

Das GG ist möglicherweise entscheidend. Die Untersuchungen an Frühchen, die mindestens 1500g zur Geburt wogen, erbrachten (bis auf die LVET1/3) durchweg höhere Kappa-Werte als die der NG mit  $GG < 1500g$ . Es zeigte sich eine wenigstens ausreichende (LPEP1/3  $\geq 1500g = 0,4$ ) bis vollkommene Wiederholbarkeit (RVET1/2 sowie RVET1/3  $\geq 1500g = 1$ ). Die Kohortenzahl ist jedoch niedrig ( $n=10$ ).

#### APGAR (siehe Tabelle 26, Seite 58)

Weder bezüglich intra-/interindividuell noch hinsichtlich PEP/VET war eine deutliche Tendenz erkennbar.

Eine besonders hohe Reliabilität erzielte die Erhebung der STIs bei FG mit einem 5-Minuten-APGAR  $\geq 8$  ( $0,556-1$ ). Dabei waren die Werte der Kinder mit einem 5-Minuten-APGAR von 5-7 regelhaft niedriger als die Werte der Kinder mit einem APGAR-Wert  $\geq 8$  (bis auf RVET1/3:  $APGAR\ 5-7 = 0,64 > APGAR \geq 8 = 0,556$ ).

#### Lebenstag (siehe Tabelle 26, Seite 58)

Bezüglich intra-/interindividueller Messungen war keine deutliche Tendenz erkennbar.

Die VET-Messungen zeigten in der Regel eine höhere Übereinstimmung als die PEP-Messungen (außer  $< 5.LT$ : LPEP1/3  $= 0,673 > \underline{LVET1/3} = 0,460$ ).

Die Kappa-Werte der Messungen der Kinder, die vor dem 5. LT untersucht wurden, waren in der Regel höher (bis auf LVET1/3:  $< 5.LT = 0,46, \geq 5.LT = 0,523$ ; RPEP1/3:  $< 5.LT = 0,276, \geq 5.LT = 0,294$ ).

Tabelle 26 Cohens Kappa Koeffizient bei gleichbleibender HF/VAS gefiltert nach SST/GG/APGAR und LT

	PEP Links		VET Links		PEP Rechts		VET Rechts	
Cohens Kappa	LPEP1/2	LPEP1/3	LVET1/2	LVET1/3	RPEP1/2	RPEP1/3	RVET1/2	RVET1/3
<b>SST&lt;224</b>	0,126 n=27	0,339 n=28	0,574 n=27	0,479 n=28	0,418 n=28	0,278 n=26	0,647 n=24	0,418 n=26
<b>SST&lt;196</b>	0,279 n=16	0,375 n=16	0,587 n=16	0,631 n=16	0,316 n=13	0,058 n=14	0,581 n=13	0,462 n=14
<b>GG&lt;1500 (g)</b>	0,158 n=24	0,379 n=25	0,504 n=24	0,491 n=25	0,288 n=21	0,228 n=23	0,577 n=21	0,33 n=23
<b>GG≥1500 (g)</b>	0,600 n=10	0,400 n=10	0,672 n=10	0,455 n=10	0,649 n=10	0,455 n=10	1 n=10	1 n=10
<b>APGAR 5-7</b>	0,095 n=19	0,081 n=19	0,486 n=19	0,466 n=19	0,254 n=18	0,242 n=18	0,640 n=18	0,640 n=18
<b>APGAR ≥8</b>	0,744 n=11	0,8 n=12	0,577 n=11	0,596 n=12	0,656 n=11	0,671 n=12	1 n=11	0,556 n=12
<b>LT &lt; 5</b>	0,355 n=16	0,673 n=17	0,716 n=16	0,460 n=17	0,527 n=13	0,276 n=15	1 n=13	0,706 n=15
<b>LT ≥ 5</b>	0,302 n=18	-0,043 n=18	0,386 n=18	0,523 n=18	0,217 n=18	0,294 n=18	0,557 n=18	0,446 n=18

LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, LPEP1/2 = Übereinstimmung Betrachtung LPEP Untersucher 1; 1. vs. 2. Messung; LPEP1/3 = Übereinstimmung Betrachtung LPEP Untersucher 1 vs. Untersucher 2; LVET1/2 = Übereinstimmung Betrachtung LVET Untersucher 1; 1. vs. 2. Messung; LVET1/3 = Übereinstimmung Betrachtung LVET Untersucher 1 vs. Untersucher 2; entsprechend für rechtsventrikuläre Werte; SST = Schwangerschaftstage; GG = Geburtsgewicht; LT = Lebenstag

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2



### Sauerstofftherapie (siehe Tabelle 27)

Die intraindividuellen STI-Messungen der NG die unter Sauerstofftherapie untersucht wurden, erzielten regelhaft höhere Kappa-Werte (siehe Tabelle 27, erste Zeile).

Die Übereinstimmung der VET-Messungen war durchgehend höher als die der PEP-Messungen.

Bei zu niedriger Anzahl an Messungen war die statistische Subgruppen-Analyse der Kinder ohne Sauerstofftherapie nicht möglich.

### Beatmung (siehe Tabelle 27)

Bezüglich der intra-/interindividuellen Messungen war keine deutliche Tendenz erkennbar.

Die Übereinstimmung der VET-Messungen war meist höher als die der PEP-Messungen (bis auf RPEP1/3: Beatmung nein = 0,584; > RVET1/3: Beatmung nein = 0,333, siehe Tabelle 27).

Die Kappa-Werte beatmeter Kinder waren in der Subgruppenanalyse bis auf wenige STI-Messungen größer (außer LVET1/2, RPEP1/3).

*Tabelle 27 Cohens Kappa Koeffizient bei gleichbleibender HF/VAS gefiltert nach O<sub>2</sub>-Gabe und Beatmung*

	PEP Links		VET Links		PEP Rechts		VET Rechts	
Cohens Kappa	LPEP1/2	LPEP1/3	LVET1/2	LVET1/3	RPEP1/2	RPEP1/3	RVET1/2	RVET1/3
<b>O<sub>2</sub> ja</b>	0,353 n=32	0,304 n=32	0,655 n=32	0,477 n=32	0,375 n=30	0,295 n=31	0,734 n=30	0,687 n=31
<b>O<sub>2</sub> nein</b>	n.m. n=2	n.m. n=3	n.m. n=2	n.m. n=3	n.m. n=1	n.m. n=2	n.m. n=1	n.m. n=2
<b>Beatmung ja</b>	0,475 n=19	0,352 n=19	0,513 n=19	0,631 n=19	0,536 n=16	-0,077 n=17	0,862 n=16	0,746 n=17
<b>Beatmung nein</b>	-0,136 n=15	0,304 n=16	0,674 n=15	0,351 n=16	0,297 n=15	0,584 n=16	0,583 n=15	0,333 n=16

*LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, LPEP1/2 = Übereinstimmung Betrachtung LPEP Untersucher 1; 1. vs. 2. Messung; LPEP1/3 = Übereinstimmung Betrachtung LPEP Untersucher 1 vs. Untersucher 2; LVET1/2 = Übereinstimmung Betrachtung LVET Untersucher 1; 1. vs. 2. Messung; LVET1/3 = Übereinstimmung Betrachtung LVET Untersucher 1 vs. Untersucher 2; entsprechend für rechtsventrikuläre Werte; O<sub>2</sub> = Sauerstoff, n.m. = statistische Analyse nicht möglich*

*Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2*

### Katecholamingabe (siehe Tabelle 28)

Die Untersuchung zeigte hinsichtlich der intraindividuellen STI-Messungen meist verlässlichere Werte (bis auf LPEP1/2: KA nein =0,091 < LPEP1/3: KA nein =0,422; RPEP1/2: KA nein =0,452 < RPEP1/3: KA nein =0,627).

Die Übereinstimmung der VET-Messungen war meist höher als die der PEP-Messungen (bis auf RPEP1/3: KA nein =0,627; > RVET1/3: KA nein =0,513).

In dem Subgruppenvergleich gab es keine klare Überlegenheit.

### PDA (siehe Tabelle 28)

Bei Kindern, die einen PDA hatten, zeigten die intraindividuellen Messungen meist höhere Kappa-Werte (bis auf RVET1/2=0,75 < RVET1/3=0,806). Bei den Kindern ohne PDA hingegen, lagen die interindividuellen Werte regelhaft höher (bis auf RVET1/2=0,756 > RVET1/3=0,432).

Die Übereinstimmung der VET-Messungen war meist höher als die der PEP-Messungen (bis auf RPEP1/3: PDA nein =0,444; > RVET1/3: PDA nein =0,432).

In dem Subgruppenvergleich gab es keine klare Überlegenheit.

*Tabelle 28 Cohens Kappa Koeffizient bei gleichbleibender HF/VAS gefiltert nach KA-Gabe und PDA*

	PEP Links		VET Links		PEP Rechts		VET Rechts	
Cohens Kappa	LPEP1/2	LPEP1/3	LVET1/2	LVET1/3	RPEP1/2	RPEP1/3	RVET1/2	RVET1/3
KA ja	0,451 n=14	0,196 n=15	0,664 n=14	0,565 n=15	0,360 n=12	-0,156 n=14	0,75 n=12	0,659 n=14
KA nein	0,091 n=20	0,422 n=20	0,508 n=20	0,450 n=20	0,452 n=19	0,627 n=19	0,756 n=19	0,513 n=19
PDA ja	0,198 n=13	0,156 n=13	0,768 n=13	0,393 n=13	0,515 n=12	0,103 n=13	0,750 n=12	0,806 n=13
PDA nein	0,403 n=21	0,450 n=22	0,437 n=21	0,580 n=22	0,219 n=19	0,444 n=20	0,756 n=19	0,432 n=20

LPEP = linksventrikuläre Präejektionszeit, RPEP = rechtsventrikuläre Präejektionszeit, LVET = linksventrikuläre Ejektionszeit, RVET = rechtsventrikuläre Ejektionszeit, LPEP1/2 = Übereinstimmung Betrachtung LPEP Untersucher 1; 1. vs. 2. Messung; LPEP1/3 = Übereinstimmung Betrachtung LPEP Untersucher 1 vs. Untersucher 2; LVET1/2 = Übereinstimmung Betrachtung LVET Untersucher 1; 1. vs. 2. Messung; LVET1/3 = Übereinstimmung Betrachtung LVET Untersucher 1 vs. Untersucher 2; entsprechend für rechtsventrikuläre Werte; KA = Katecholamine; PDA = persistierender Ductus arteriosus

Anzahl „n“ der einfließenden Untersuchungen individuell angegeben, fehlende Untersuchungen siehe 4.2

## 5.8 Zusammenfassung der wichtigsten Ergebnisse

In den Bland et Altmann Diagrammen der STIs zeigte sich, dass die Mehrheit der Ausreißer FG waren. Ein weiteres häufig assoziiertes Charakteristikum der aberranten Werte in der Bland et Altmann Analyse war ein GG<1500 g. Bei Betrachtung sämtlicher STIs lagen immer >90% der Differenzen innerhalb des jeweiligen 95%-Konfidenzintervalls.

Bezüglich der Repeatability Indices (RI) konnte zunächst gezeigt werden, dass nur die intraindividuell wiederholten VET-Messungen im Akzeptanzbereich  $RI \leq 0,1$  lagen. Nachdem die Kohorte bzgl. gleichbleibender HF ( $[HF1-HF2] \leq 0,05 \cdot HF1$ ) und unveränderter Vigilanz ( $VAS1-VAS2 \leq 10\%$ ) selektiert wurde, war auch die interindividuelle LVET-Messung ausreichend reproduzierbar.

Bei weiterer Analyse der akzeptablen RIs hinsichtlich möglicher Einflussfaktoren, wurden die schlechtesten RIs zumeist in der Kategorie Geburtsgewicht (Subgruppe GG <1500g) bzw. APGAR (Subgruppe 5-Minuten APGAR von 5-7) detektiert. Kinder mit einem GG  $\geq 1500$ g oder einem APGAR  $\geq 8$  erzielten, innerhalb der akzeptablen RIs, zufriedenstellende Werte.

Die anschließende statistische Auswertung der Daten anhand des Kappa-Koeffizienten nach Cohen demonstrierte ebenfalls eine ausgeprägte Reliabilität der intraindividuellen VET-Messungen, sowohl in der Gesamtheit aller Untersuchungen, als auch selektiert nach gleichbleibender HF ( $[HF1-HF2] \leq 0,05 \cdot HF1$ ) und unveränderter Vigilanz ( $VAS1-VAS2 \leq 10\%$ ).

Auch die Kappa Analyse zeigte eine besonders hohe Übereinstimmung der wiederholten STI-Messungen bei Probanden mit einem APGAR-Wert  $\geq 8$  und einem GG  $\geq 1500$ g.

Insgesamt muss jedoch gesagt werden, dass keine klaren Tendenzen in den nach Kategorien gefilterten Daten feststellbar waren. Weder bei der RI-, noch bei der Kappa-Analyse waren klare Trends in den Subgruppenanalysen erkennbar. Es gab vereinzelt Hinweise auf einen möglichen Einfluss der einzelnen Charakteristika. Wie bereits erwähnt, waren bei der RI-Ermittlung in den Kategorien Gewicht und APGAR besonders viele Maxima (GG<1500g, APGAR 5-7) und Minima (GG $\geq 1500$ g, APGAR  $\geq 8$ ) der akzeptablen RIs ( $RI < 0,1$ ) angesiedelt. In der Kappa Analyse konnten diese leichten Tendenzen erneut aufgezeigt werden. Auch die Ausreißer der Bland et. Altmann Plots waren häufig FG mit GG<1500g.

## 6. Diskussion

Um die Wiederholbarkeit und Zuverlässigkeit der Erhebung der STIs zu analysieren, wurden in der vorliegenden retrospektiven Studie Daten des Zeitraums 2016 bis 2018 ausgewertet. Die dabei ermittelten STIs entsprechen den Normwerten der Literatur.<sup>15,42,44</sup> Die RVETs waren länger als die LVETs, während bei den PEPs kein Unterschied beobachtet wurde (siehe Tabelle 7, Seite 28). Hinsichtlich der Normbereiche der STIs einzelner Untergruppen ist in der Literatur beschrieben, dass VETs bei erwachsenen Frauen signifikant länger sind.<sup>2</sup> Für NG wurde diesbezüglich unseres Wissens nach bislang kein geschlechtsspezifischer Unterschied beschrieben.<sup>44</sup> In der hier untersuchten Kohorte konnte ebenfalls keine derartige Abweichung ermittelt werden.

Angeregt durch die Ergebnisse aktueller Publikationen, beispielsweise von Schwarz, Preusche et al. 2016 war das Bestreben dieser Arbeit, die Reliabilität der STIs anhand einer größeren Kohorte und differenziert nach möglichen Einflussfaktoren zu überprüfen. Die wichtigste Erkenntnis für den klinischen Alltag ist, dass nur die dopplersonographische VET-Ermittlung eine ausreichende Reproduzierbarkeit und damit diagnostische Verlässlichkeit gewährleistet. Die PEP hat im optimalen Fall eine moderate Verlässlichkeit geboten. Zudem zeigten intraindividuell wiederholte Messungen in der Regel eine höhere Übereinstimmung. Auch ein 5-Minuten-APGAR  $\geq 8$ , sowie ein Geburtsgewicht  $\geq 1500\text{g}$  scheinen die Übereinstimmung der Messungen positiv zu beeinflussen. Umgekehrt verringern ein 5-Minuten-APGAR  $< 8$  und ein GG  $< 1500\text{g}$ , sowie Frühgeburtlichkeit die Genauigkeit.

Unseres Wissens nach, ist dies die erste Arbeit, die sich derart spezifisch, nicht nur mit der intra-, sondern auch der interindividuellen Übereinstimmung und Reproduzierbarkeit wiederholter STI-Messungen auseinandersetzt. Die statistische Auswertung erfolgte zunächst anhand des RI, anschließend wurden die Daten erneut mit Hilfe des Cohen's Kappa Koeffizienten hinsichtlich ihrer klinischen Verlässlichkeit untersucht. Die Studie ermöglicht folglich nicht nur eine quantitative, sondern auch eine qualitative Beurteilung hinsichtlich der Zuverlässigkeit der Untersuchungen. Der umfangreiche Datensatz mit 98 Teiluntersuchungen zu je drei Messungen an insgesamt 75 Probanden umfasst, insbesondere für den Teilbereich der Neonatologie, Datenmaterial einer außerordentlich großen Kohorte. Die vergleichsweise hohe Stichprobenzahl gewährleistet somit eine relativ ausgeprägte statistische Aussagekraft.

Weiterhin wurde versucht, Faktoren, die möglicherweise die Reproduzierbarkeit der Datengenerierung beeinflussen (Diskrepanz der HF>5%; Abweichung >10% auf der VAS zwischen den Messungen), auszuschließen. Die erneute Betrachtung der Untersuchungen, die diese Voraussetzungen erfüllten, gewährleistete somit eine zusätzliche Optimierung der Aussagekraft der erhobenen Ergebnisse. Die Tatsache, dass im klinischen Alltag eine kurzfristige Veränderung der Vigilanz der Kinder (gemessen mit Vergleich von HF/VAS) zwischen zwei Messungen nicht auszuschließen ist, wurde bedacht.

Auch der standardisierte Untersuchungsablauf und die spezifische Aufteilung der Kohorte anhand diverser Charakteristika, beispielsweise der Untersucherexpertise, sorgen für eine hohe Transparenz bezüglich der Einflussfaktoren auf die Wiederholbarkeit der Messungen.

Kritisch zu bewerten ist, dass wir zur Eruierung der Wiederholbarkeit der Daten anhand des RI, aufgrund der Ermangelung eines in der Literatur beschriebenen Grenzwertes, nach Rücksprache mit dem IMISE (Institut für Medizinische Informatik, Statistik und Epidemiologie) einen zentrumsinternen Akzeptanzbereich von 10% festgelegt haben (entspricht einem  $RI \leq 0,1$ ). Dieser ist eng gefasst, wenn man bedenkt, dass wir Differenzen von Zeitintervallen im Millisekundenbereich untersuchen. Bezüglich der Kappa-Analyse gilt es zu beachten, dass durch die Auswertung kategorialer Daten (über/im/unter dem Normbereich) Information verloren geht.

Die Probandenzahl ist insgesamt relativ hoch für eine monozentrische Studie des Bereiches Neonatologie, trotzdem ist zu erwähnen, dass die untersuchten Subkohorten teilweise nur wenige Kinder einbeziehen. Ein weiterer Kritikpunkt ist, dass es unvollständig ausgefüllte Untersuchungsbögen gab, beispielsweise aufgrund erschwerter Messung bestimmter STIs bei schlechten Schallbedingungen. In unserer statistischen Auswertung wurde stets versucht, das Maximum an Daten auszuwerten, mit entsprechender Kennzeichnung abweichender Kohortenzahlen aufgrund fehlender Messwerte. Die notwendige Transparenz ist damit gegeben.

Auch wenn es noch keine prospektiv-kontrollierte Anschlussuntersuchung der Ergebnisse gab, konnten durch die Studie wichtige Hypothesen für den Umgang mit STIs in der Klinikroutine aufgestellt werden. Bisher publizierte Studien umfassten nur recht kleine Fallzahlen, eine wirklich umfangreiche, kritische statistische Auseinandersetzung hinsichtlich

der Anwendung der STIs in der Neonatologie ist uns nicht bekannt. Dies ist die erste Studie, die sich derart detailliert und spezifisch mit der Reproduzierbarkeit und Verlässlichkeit der STI-Messungen befasst.

In der Literatur wurde die echokardiographische Messung der STIs wiederholt als technisch leicht zu erlernende Methode beschrieben, die sich durch ihre Nicht-Invasivität und leichte Durchführbarkeit auszeichnet.<sup>1,2,14</sup> Eine hohe Verlässlichkeit der STI-Messungen ist nicht nur für Erwachsene, sondern auch Kinder und NG angegeben.<sup>15,45,46</sup> Die Technik wird hinsichtlich ihrer Aussagekraft bezüglich der globalen kardialen Funktionalität als sensitiv, präzise und wiederholbar bezeichnet.<sup>23,47,48</sup> Demnach könnten auch bei schwer kranken Kindern STI-Messungen durchgeführt werden. Invasive Methoden die dem Monitoring der Auswurfleistung des Herzen eines NG dienen, sind häufig durch die geringe Körpergröße und dem damit verbundenen geringen Blutvolumen, sowie einer noch instabilen Hämodynamik der NG- und FG, nur limitiert einsetzbar.<sup>17,49</sup>

Echokardiographische Mehrfachmessungen, auch im Verlauf, sind leicht zu realisieren und würden die erweiterte Nutzung der STIs als Verlaufsparemeter ermöglichen, auch um beispielsweise den Effekt einer neuen Therapie zu beobachten.<sup>23</sup>

Die Literatur ist jedoch häufig veraltet und wurde lange nicht hinterfragt. Eine statistische Auseinandersetzung erfolgte in nur wenigen Studien.

In aktuelleren Studien wurde eine schlechte Reproduzierbarkeit der STI-Messung im Rahmen der PDA Diagnostik bei NG beschrieben.<sup>31</sup> Auch die vorliegende Arbeit demonstriert im Gesamtbild nur eine moderate Verlässlichkeit der Erhebung STI in der kardialen Beurteilung NG.

Insbesondere die PEP-Messung, wird in der Literatur mehrfach als einfache Methode mit einer hohen Reproduzierbarkeit beschrieben. Dabei wird die Reproduzierbarkeit in besagter Literatur jedoch nicht statistisch dargelegt.<sup>15,21</sup> In der hier präsentierten Studie zeigt die PEP weder inter-, noch intraindividuell unabhängig von der statistischen Analyse eine befriedigende Wiederholbarkeit.

Die PEP sollte folglich nicht als Standarddiagnostikum zur Beurteilung der kardialen Kontraktilität Neugeborener dienen. Die VET-Messungen zeigte hingegen eine gute

Reproduzierbarkeit. Insbesondere die Werte der intraindividuell wiederholten Messungen zeigten eine hohe Übereinstimmung. Die dopplersonographische Methode zur VET-Ermittlung sollte nicht zuletzt wegen ihrer leichten Durchführbarkeit und ubiquitären Verfügbarkeit zentrumsübergreifend eine tragende Rolle bei der alltäglichen kardialen Beurteilung Neugeborener spielen. Sie kann insbesondere zur Detektion eines hsPDA beitragen.

Eine ältere Studie von Van Dijk et al. aus dem Jahr 1996 beschreibt ebenfalls eine unzureichende Verlässlichkeit der PEP. Im Rahmen der Studie wurden rv STI-Messungen an 44 Patienten (0,1-16,2 J.) erhoben und die inter- und intraindividuelle Variabilität anhand der SD untersucht. Hierbei wurde die PEP-Messung (SD der absoluten Differenzen der interindividuell wiederholten PEP-Messungen=15,6%; SD der absoluten Differenzen der intraindividuell wiederholten PEP-Messungen=11,2%) nur als ausreichend reproduzierbar beschrieben. Aus den Daten lässt sich außerdem sowohl der intra-, als auch der interindividuelle RI für die rv STI-Messungen errechnen. Die Werte liegen sowohl inter-, als auch intraindividuell schlechter als unsere Ergebnisse und wären nach unseren Maßgaben nicht akzeptabel (intraindividuell:  $RI-RPEP_{1/2}=0,29$ ; interindividuell:  $RI-RPEP_{1/3}=0,53$ ; Vgl. mit den Werten der vorliegenden Studie siehe Tabelle 9, Seite 33, erste Zeile). Anders zeigt sich die Variabilität der repetitiven RVET-Messung. Mit einer SD der absoluten Differenzen der interindividuell wiederholten Messungen von 8.3% und einer SD der absoluten Differenzen der intraindividuell wiederholten Messungen von 6% wird die Erhebung der RVET als gut reproduzierbare Methode eingeschätzt. Die RIs für die RVET-Messungen glichen noch am ehesten den Ergebnissen unserer Studie (intraindividuell: 0,125; interindividuell: 0,18; Vgl. mit den Werten der vorliegenden Studie siehe Tabelle 9, Seite 33, Zeile 1).<sup>50</sup>

Andere Studien, die die Varianzkoeffizienten bestimmten, um die Wiederholbarkeit der dopplersonographischen STI-Messung zu beurteilen, kamen zu dem Urteil, dass nicht nur die VET-, sondern auch die PEP-Messungen intra- und interindividuell durchaus verlässlich seien (intraindividueller Variationskoeffizient  $PEP=1,7\%$ ,  $LVET=1,7\%$ ,  $PEP-LVET\text{-ratio}=2,9\%$ ). Interindividuell lagen die Werte etwas schlechter ( $PEP=2,6\%$ ,  $LVET=2,0\%$ ,  $PEP-LVET\text{-ratio}=3,9\%$ ).<sup>45</sup> Es ist aus unserer Sicht allerdings sehr fraglich, ob der Varianzkoeffizient für eine derartige Beurteilung der richtige Parameter ist. Die von uns vorgenommene statistische Betrachtung mit Bland-Altman-Plots, Berechnung von RI und RC, sowie die qualitative

Cohens-Kappa-Analyse ist dafür eine viel kritischere (und damit genauere) Methode. Weiterhin untersuchten wir verschiedene mögliche Faktoren, die einen mutmaßlichen Einfluss auf die Reproduzierbarkeit der Messungen haben könnten. Dazu zählten wir das Erfahrungslevel des Untersuchers, das Geschlecht der Probanden, das Gestationsalter, das Geburtsgewicht, den 5-Minuten-APGAR-Wert, eine zusätzliche O<sub>2</sub>-Vorlage, Beatmungstherapie, Katecholamintherapie und das Vorhandensein eines PDA.

Dabei kamen wir zu dem Ergebnis, dass ein 5-Minuten-APGAR-Wert  $\geq 8$  und ein GG  $\geq 1500$ g einen positiven Einfluss hinsichtlich der Übereinstimmung und Verlässlichkeit der Messungen an Neugeborenen zu haben scheinen. Umgekehrt zeigten wiederholte Messungen an Neugeborenen mit einem niedrigen APGAR oder GG weniger Übereinstimmung. Demnach sollte die VET-Messung insbesondere für Kinder mit einem 5-Minuten-APGAR-Wert  $\geq 8$  und einem GG  $\geq 1500$ g in Frage kommen. Es wurde hierbei nicht untersucht wie die multifaktoriellen Zusammenhänge zwischen diesen beiden Messgrößen (Geburtsgewicht und APGAR) sind. Es ist bekannt, dass der APGAR mit dem Gestationsalter (und damit auch mit dem Geburtsgewicht) positiv korreliert. Deshalb ist eine Ko-Abhängigkeit dieser Werte wahrscheinlich.

Wichtig ist allerdings, dass die Wiederholbarkeit, vor allem der interindividuellen Ergebnisse, sich bei strikter Einhaltung gleicher Untersuchungsbedingungen hinsichtlich VAS und HF bessert. Dieses Ergebnis kann aus theoretischen Gründen postuliert werden, weil die Werte für die LVET und die RVET von der Herzfrequenz abhängen. Die Daten konnten statistisch beweisen, dass die Wiederholbarkeit der Ergebnisse in der Mehrheit erwartungsgemäß besser ist, wenn die Vigilanz des untersuchten Kindes gleichbleibt.

In einer Studie von 2016 <sup>31</sup> wurden 19 FG mit einem GG  $\leq 1500$ g und klinischen Hinweisen auf einen PDA untersucht. Die interindividuelle Reproduzierbarkeit verschiedener gängiger Methoden zur Beurteilung der kardialen Funktionalität Neugeborener wurde analysiert. Ermittelt wurden die RIs der dopplersonographischen Flussmessung in der Arteria cerebri anterior, der Flussmessung im Truncus coeliacus, der LA/Ao-ratio und der lv WR (LPEP/LVET). Dabei lagen die RI-Werte für die Flussmessung in der Arteria cerebri anterior bei 0,14 und für die Flussmessung im Truncus coeliacus bei 0,11. Diese Werte waren besser als die Messungen der LA/Ao-ratio (RI=0,16) und der lv WR (RI=0,23).<sup>31</sup>



Keines dieser Ergebnisse hätte dabei unseren Anforderungen ( $RI < 0,1$ ) bezüglich einer akzeptablen Wiederholbarkeit entsprochen. Das Geburtsgewicht  $\leq 1500\text{g}$  kann bei Betrachtung unserer Studie ein Grund für den hohen RI-Wert der lv WR sein.

In der vorliegenden Arbeit wurden die PEP und die VET separat, und nicht die WR (PEP/LVET), analysiert. Dies erschien uns genauer, um diskrepante Ergebnisse genau aufdecken zu können. Die WR als sekundär errechneter Wert hat gegenüber den Originalwerten naturgemäß zahlreiche Unwägbarkeiten.

Aufgrund letztgenannter Studie<sup>31</sup> kam nun die Frage auf, ob die STI-Messungen mit aktuell verfügbaren Alternativmethoden konkurrieren können. In der Literatur gibt es entsprechend vergleichende Studien.<sup>16,22,48,51</sup>

Eine multizentrische Analyse vergleicht die Wiederholbarkeit der STI-Messungen mit neueren, jedoch schwieriger zu erlernenden und hochtechnisierten echokardiographischen Methoden, wie dem Gewebsdoppler und der Speckle-tracking-strain-Analyse. Diese Untersuchungen sind mit einem hohen Zeitaufwand verbunden. Die intraindividuelle Reproduzierbarkeit der dopplersonographischen STI-Erfassung war sowohl für die PEP (Varianz=2%), als auch für die LVET (Varianz=4%) besser als die echokardiographische Ermittlung der LVEF (linksventrikuläre Ejektionsfraktion, Varianz=13%) und die Global-longitudinal-strain-Analyse ([GLS-Analyse]; Varianz=7%). Die Ergebnisse hinsichtlich der interindividuellen Wiederholbarkeit waren ebenfalls insbesondere für die STIs zufriedenstellend (Varianz: LPEP 3%; LVET 4,5%; LVEF 15%; GLS 8%). Untersucht wurden 177 Erwachsene, davon 134 Herzinsuffizienzpatienten.<sup>16</sup> Wir möchten an dieser Stelle noch einmal betonen, dass wir für eine Fragestellung hinsichtlich der Reproduzierbarkeit mehrerer Messungen eine Auswertung mit den von uns angewendeten Methoden favorisieren. Auch wurden in der genannten Studie für die Analyse 10 repetitive Messungen durchgeführt und anschließend die Varianz bestimmt, ein Vorgehen, was in der Neonatologie kaum möglich erscheint.

Weitere nichtinvasive Messmethoden abseits der Echokardiographie wurden bereits hinsichtlich ihrer Exaktheit, vergleichend zur Echokardiographie als Referenzmethode, untersucht. Dabei wird der STI-Ermittlung mit Hilfe der Seismokardiographie (SCG)<sup>48</sup>, aber auch der Phonokardiographie (PCG)<sup>51</sup> eine akzeptable Vergleichbarkeit zur Referenzmethode

zugesprochen. Die Empfehlungen zur Anwendung der Photoplethysmographie sind bei einem Datensatz von  $n=13$  statistisch wenig aussagekräftig.<sup>48</sup> Die Impedanz-Kardiographie (ICG) hingegen zeigte regelhaft Messabweichungen gegenüber den echokardiographisch ermittelten Werten. Die PEP und auch die VET wurden meist überschätzt.<sup>48,51</sup>

Carvalho et al. haben beispielsweise die Genauigkeit der Ergebnisse hinsichtlich der Erhebung von STIs zwischen verschiedenen nicht-invasiven Methoden mit dem aktuellen Goldstandard, der Echokardiographie, verglichen. Das Ziel der Studie war, das optimale Verfahren zu ermitteln, mit dem Ansatz, zukünftige Long-term-follow-Up-tools für die globale Herzfunktion von erwachsenen Patienten zu entwickeln. Dabei konnte festgestellt werden, dass die PEP und auch die LVET analog zur o.g. Studie regelhaft durch die ICG überschätzt wurden. Die Anwendung der ICG zeigte die schlechtesten Ergebnisse, wohingegen der Gebrauch der Photoplethysmographie (PPG) oder die Nutzung eines PCG potentielle Monitoring-Techniken für die Zukunft darstellen.<sup>52</sup>

Bezüglich der STI-Erhebung mittels PPG gibt es außerdem eine Studie, die die Messgenauigkeit der Methode vergleichend zur Beurteilung verschiedener echokardiographischer Parameter untersucht hat. Die Analyse erfolgte im Rahmen der Bewertung des hämodynamischen Status erwachsener Patienten mit DDD-Schrittmacher nach kardialer Resynchronisationstherapie. Hierbei konnte gezeigt werden, dass der Variationskoeffizient (VC) des Systolischen Index ( $[SI=LVET/PEP]$ ;  $VC=3,4\%$ ) im Vergleich zum systolische Blutdruck (SBP;  $VC=5,7\%$ ) und zum dopplersonographisch ermittelten aortic-velocity-time-integral (AoVTI;  $VC=7,8\%$ ) geringer war, was für geringere Messfehler spricht.<sup>22</sup> Der SI-Messung mittels PPG wurde hierbei eine hohe diagnostische Genauigkeit und Sensitivität zugesprochen. Eine differenzierte Untersuchung der PEP und VET fand diesbezüglich nicht statt.

Die beschriebenen Verfahren stellen somit größtenteils (bis auf die ICG) mögliche Alternativmethoden zur echokardiographischen Ermittlung systolischer Zeitintervalle dar. Die Messergebnisse variierten kaum gegenüber der Referenzmethode. Allerdings sind die Diagnostiktools nicht ubiquitär verfügbar, zeitlich und technisch aufwendiger und es wurde in

den beschriebenen Studien nur ein Vergleich zur Referenzmethode, der Echokardiographie, gezogen.

Unsere Daten konnten zeigen, dass die dopplersonographischen intraindividuellen VET-Messungen, und unter gleichbleibenden Voraussetzungen bezüglich der VAS und der HF auch die interindividuellen LVET-Messungen, reproduzierbar sind. Dies gilt insbesondere für Kinder  $\geq 1500\text{g}$  und für Kinder mit einem APGAR von  $\geq 8$ . Weitere Auffälligkeiten der Messwertwiederholbarkeit hinsichtlich bestimmter Voraussetzungen (Geschlecht, Erfahrung der Untersucher etc.) waren nicht festzustellen. Entsprechend kann die Methode als Standarddiagnostikum im neonatologischen Alltag zur Beurteilung des hämodynamischen Status Neugeborener in Betracht gezogen werden. Derzeit gibt es dahingehend noch keine generellen Empfehlungen.

Die Methode ist simpel, kostensparend, leicht zu erlernen, ubiquitär verfügbar und nicht invasiv. Verlaufsbeurteilungen sind möglich. Der Messung systolischer Zeitintervalle wird in der aktuellen Literatur vergleichend zu anderen echokardiographisch eruierbaren Parametern teilweise eine bessere (intra-/interindividuelle Reproduzierbarkeit vgl. Varianzanalyse zu LVEF, GLS<sup>16</sup>) , teilweise eine schlechtere (interindividuelle Reproduzierbarkeit vgl. Resistance Index zu Flussmessung ACA, CA und LA/Ao-ratio<sup>31</sup>) Reproduzierbarkeit zugesprochen. Zu beachten ist dabei, dass die vergleichende Studie zur LVEF/GLS erwachsene Probanden untersucht hat und die letztgenannte Studie nur 16 Neugeborene eingeschlossen hat. Damit ist die ausreichende Aussagekraft der Daten in gewisser Weise eingeschränkt.

Auch bei selektiven Untersuchungen zu rechtsventrikulären STIs sind dabei sowohl intra-, als auch interindividuell (zumindest hinsichtlich der VET) grenzwertig akzeptable Werte bezüglich der Wiederholbarkeit der Messungen beschrieben.<sup>50</sup> Es gibt vergleichende Analysen zu Alternativmethoden gegenüber der echokardiographischen Erhebung von STIs. Die PPG/PCG/SCG stellen dabei andere verlässliche Methoden dar, um STI-Messungen durchzuführen. Die ICG hingegen scheint weniger verlässlich zu sein.<sup>22,48,51,52</sup> Die Beurteilung erfolgte jedoch nur anhand einer Varianzanalyse. Diese Methoden sind außerdem technisch aufwendiger, schwerer zu erlernen und nicht ubiquitär verfügbar.

In der aktuellen Literatur fehlen vergleichende Studien zur inter-/intraindividuellen Reproduzierbarkeit der Ermittlung von STIs anhand besagter alternativer nichtinvasiven Messmethoden (ICG, PCG, PPG, SCG) in Gegenüberstellung zur dopplersonographischen STI-Ermittlung bei NG. Dies könnte Inhalt zukünftiger Untersuchungen sein. Auch gilt es, die STI-Erhebung mittels weiterer echokardiographischer Messverfahren hinsichtlich ihrer Wiederholbarkeit zu überprüfen. Beispielsweise fehlen detaillierte statistische Analysen für STI-Untersuchungen mit der Strain-Analyse mit Speckle-Tracking oder mit Hilfe der Tissue-Doppler-imaging-Methode, die ebenfalls in den letzten Jahren zur STI-Ermittlung genutzt wurden.<sup>48</sup>

Nicht betrachtet wurden, wegen des Informationsverlustes gegenüber der Einzelbetrachtung von PEP und LVET, die RIs und der Kappa Koeffizient der WR (PEP/LVET). In zukünftigen Studien könnte dies dennoch ein Untersuchungsschwerpunkt sein.

Schwierigkeiten bestehen bezüglich der Definition eines erfahrenen Untersuchers. Dahingehend wären zentrumsinterne oder sogar zentrumsübergreifende regelmäßige standardisierte Weiterbildungen wünschenswert. Auch könnte eine weitere Studie mit Reifgeborenen erfolgen. Möglicherweise wäre hierbei eine weitere Steigerung der Reproduzierbarkeit der Messungen aufgrund besserer Schallbedingungen, längerer STIs etc. möglich.

Langfristig wird in größeren Studien untersucht werden müssen, ob sich die STI-Erhebung gegenüber neueren Methoden, wie beispielsweise der Speckle-Tracking-Echokardiographie, durchsetzen kann.

Die dopplersonographische Beurteilung der STIs war eine der ersten nichtinvasiven Methoden, um die globale kardiale Funktion bewerten zu können. Die einst zugeschriebene Verlässlichkeit der STI-Messungen wurde lange Zeit nicht hinterfragt. In den letzten Jahren wurden jedoch Studien publiziert, die die klinische Anwendbarkeit und Relevanz aufgrund mangelnder Reproduzierbarkeit und Übereinstimmung intra- und interindividueller Messungen angezweifelt haben. Die vorliegende Studie hat dahingehend zu wichtigen Erkenntnissen bezüglich der klinischen Anwendbarkeit der STI-Erhebung geführt.

## 7. Zusammenfassung

Dissertation zur Erlangung des  
akademischen Grades:

*Doctor Medicinae (Dr. med.)*

Titel:

„Qualitätsbeurteilung der Messung systolischer Zeitintervalle  
in der neonatologischen Praxis durch Analyse der intra- und  
interindividuellen Reliabilität“

Eingereicht von:

*Riekje Isberner, geboren 14.09.1994 Berlin*

Angefertigt in

*Der Klinik und Poliklinik für Kinder- und Jugendmedizin,  
Abteilung für Neonatologie*

*An der Universität Leipzig*

Betreut von:

*Prof. Dr. med. Matthias Knüpfer*

*Facharzt für Kinder- und Jugendmedizin, Neonatologie*

Einreichung:

Eine der häufigsten diagnostischen und therapeutischen Herausforderungen der Neonatologie stellen kardiovaskuläre Pathologien und im Rahmen dessen Kreislaufdysregulationen bei Neugeborenen dar. Eine adäquate postnatale Kreislaufadaptation ist entscheidend für das Überleben der Kinder.<sup>3</sup> Die Echokardiographie ist dabei ein außerordentlich bedeutendes diagnostisches Mittel zur Beurteilung der hämodynamischen Situation der Neonaten. Über die beste diagnostische Methode gibt es bisher keinen Konsens. In der vorliegenden Arbeit wurde die Wiederholbarkeit (repeatability) und die Verlässlichkeit (reliability) der dopplersonographischen Messung systolischer Zeitintervalle analysiert. Die Messung der STIs ist einfach. Sie erfordert neben der dopplersonographischen Erhebung der Flussprofile der Semilunarklappen lediglich eine parallele EKG-Ableitung. Die Präejektionsphase (PEP) umfasst das Zeitintervall ab der Ventrikeldepolarisation bis zum Beginn der Austreibungsphase. Die linksventrikuläre Ejektionsphase (LVET) ist begrenzt durch Öffnung und Schluss der Aortenklappe und die rechtsventrikuläre Ejektionsphase (RVET) entsprechend durch das Öffnen und Schließen der Pulmonalklappe.<sup>2,15,16</sup>

Die STI-Messung stellt somit eine nicht-invasive Möglichkeit dar, die globale Herzfunktion zu beurteilen<sup>1,2,14,19</sup>, weshalb die Methode deutschlandweit an mehreren Universitätskliniken routinemäßig angewandt wird. Aktuelle Studien haben die Reproduzierbarkeit der STIs hinterfragt, ausführliche statistische Analysen dahingehend gibt es nicht. Das vorrangige Bestreben dieser Arbeit war, die Reliabilität der STIs anhand einer aussagekräftigen Kohorte und differenziert nach möglichen Einflussfaktoren zu überprüfen. Unseres Wissens nach, ist dies die erste Arbeit, die sich derart spezifisch, nicht nur mit der intra-, sondern auch der interindividuellen Übereinstimmung und Reproduzierbarkeit wiederholter STI-Messungen auseinandersetzt. Die statistische Auswertung erfolgte mittels Bland et. Altmann plots, mit Hilfe des Repeatability Koeffizienten (Wiederholbarkeitskoeffizient), des Repeatability Index (Wiederholbarkeitsindex) und dem Kappa Koeffizienten nach Cohen.

Im Zeitraum vom 06.10.2016 bis zum 09.02.2018 wurden in der neonatologischen Abteilung der Universitätsklinik Leipzig alle Kinder, die in der klinischen Arbeit routinemäßig oder indikationsbezogen eine Echokardiographie erhielten, als mögliche Teilnehmer der Studie erfasst. So konnte ein umfangreicher Datensatz mit 98 Teiluntersuchungen zu je drei Messungen an insgesamt 75 Probanden erhoben werden. Die Möglichkeit sowohl einer doppelten Messung durch den ersten Untersucher, als auch einer zeitnahen Kontrollmessung (verblindet) durch einen zweiten Untersucher innerhalb einer maximalen Untersuchungsdauer von 10 Minuten war dabei Voraussetzung. Ein probandenbezogener Studienabbruch aufgrund unerwarteter Ereignisse, wie starker Unruhe, ersichtlichen Schmerzen oder vermutetem Unwohlsein, trat nicht ein. Auf gleichbleibende Untersuchungsvoraussetzungen wurde strikt geachtet. Die Dokumentation fand auf zuvor ausgearbeiteten Bögen statt. Neben den STIs wurden sämtliche unseres Erachtens nach entscheidenden Einflussfaktoren festgehalten, insbesondere die Agitiertheit der Patienten während der Untersuchung, sowie die Herzfrequenz, aber auch Charakteristika wie das Vorhandensein eines PDA, Katecholamin-, oder Beatmungstherapie, Gestationsalter, Geburtsgewicht u.v.m. Diese Faktoren wurden vor Studienbeginn anhand klinischer Erfahrungen festgelegt.

Die Arbeit demonstriert die Ergebnisse einer höchst differenzierten statistischen Auswertung des Datensatzes. Zunächst erfolgte die Generierung von Bland et. Altmann Diagrammen zur Veranschaulichung der Schwankungsbreite der Differenzen zwischen den intraindividuell/

interindividuell wiederholten Messwerten der STIs. Systematische Messfehler konnten dadurch ausgeschlossen werden. Anschließend erfolgte eine erste explizite Analyse der Reproduzierbarkeit mit Hilfe des Repeatability Coefficient (RC). Er beschreibt den Erwartungswert, unter dem mit einer Wahrscheinlichkeit von 95% die absolute Differenz zwischen zwei Ergebnissen einer wiederholten STI-Analyse liegt.<sup>40</sup> Der Repeatability index (RI) ist das Ergebnis der Division des RC durch den Mittelwert aller Messungen. Er stellt somit das Verhältnis der doppelten Standardabweichung des Mittelwertes der Differenzen sämtlicher Messungen gegenüber dem Mittelwert der Ergebnisse aller Messungen dar.<sup>31</sup> Außerdem analysierten wir die Daten durch den Koeffizienten nach Kappa qualitativ. Der Wert beschreibt das Verhältnis zwischen der zufällig zu erwartenden Übereinstimmung  $p_e$  und der tatsächlich beobachteten Übereinstimmung.<sup>41</sup>

Es erfolgte in jeder Untersuchung zunächst die Analyse des gesamten Datensatzes, anschließend gefiltert nach entsprechenden Kategorien hinsichtlich festgelegter möglicher Einflussfaktoren. Bei der RC-, RI-, sowie der Cohens Kappa-Analyse filterten wir die Daten in einer zusätzlichen Auswertung nach gleichbleibenden Untersuchungsvoraussetzungen bezüglich Agitiertheit und Herzfrequenz des Neugeborenen während der Untersuchung.

Die Mehrheit der Ausreißer in den Bland et. Altmann Diagrammen waren Frühgeborene. Ein weiteres häufig assoziiertes Charakteristikum der aberranten Werte in der Bland et Altmann Analyse war ein GG < 1500 g. Bei sämtlichen STI-Betrachtungen lagen >90% der Differenzen innerhalb des jeweiligen 95%-Konfidenzintervalls. Bezüglich der Repeatability Indices (RI) konnte zunächst gezeigt werden, dass nur die intraindividuell wiederholten VET-Messungen im Akzeptanzbereich  $RI \leq 0,1$  lagen. Dieser Wert wurde bei Ermangelung eines offiziellen Richtwertes in der Literatur durch uns festgelegt. Nachdem die Kohorte bzgl. gleichbleibender HF und unveränderter Vigilanz selektiert wurde, war auch die interindividuelle LVET-Messung ausreichend reproduzierbar. Bei der anschließenden Analyse der akzeptablen RIs hinsichtlich möglicher Einflussfaktoren, wurden die schlechtesten RIs zumeist in der Kategorie Geburtsgewicht (Subgruppe GG < 1500g) bzw. APGAR (Subgruppe 5-Minuten APGAR von 5-7) detektiert. Kinder mit einem GG  $\geq 1500$ g oder einem APGAR  $\geq 8$  erzielten innerhalb der akzeptablen RIs äußerst zufriedenstellende Werte. Die statistische Auswertung der Daten anhand des Kappa-Koeffizienten nach Cohen zeigte ebenfalls eine ausgeprägte Reliabilität der intraindividuellen VET-Messungen. Sowohl in der Gesamtheit

aller Untersuchungen, als auch selektiert nach gleichbleibender HF und Vigilanz. Auch die Kappa Werte wiesen auf eine besonders hohe Reliabilität der STI-Messungen bei Probanden mit einem APGAR-Wert  $\geq 8$  und einem GG  $\geq 1500$ g hin.

Insgesamt muss jedoch gesagt werden, dass keine klaren Tendenzen in den nach Kategorien gefilterten Daten feststellbar waren. Weder bei der RI-, noch bei der Kappa-Analyse waren klare Trends in den Subgruppenanalysen erkennbar. Es gab vereinzelt Hinweise auf einen möglichen Einfluss der einzelnen Charakteristika.

Die untersuchte Methode ist simpel, kostensparend, leicht zu erlernen, ubiquitär verfügbar und nicht invasiv. Verlaufsbeurteilungen sind möglich. Der Messung systolischer Zeitintervalle wird in der aktuellen Literatur vergleichend zu anderen echokardiographisch eruierbaren Parametern teilweise eine bessere (vgl. Varianzanalyse zu LVEF, GLS),<sup>16</sup> teilweise eine schlechtere Reproduzierbarkeit zugesprochen (vgl. RI-Erhebung zu Flussmessung ACA, CA und LA/Ao-ratio).<sup>31</sup>

In der aktuellen Literatur fehlen vergleichende Studien zur inter-/intraindividuellen Reproduzierbarkeit der Ermittlung von STIs anhand besagter alternativer nichtinvasiven Messmethoden (ICG, PCG, PPG, SCG) in Gegenüberstellung zur dopplersonographischen STI-Ermittlung bei NG. Dies könnte Inhalt zukünftiger Untersuchungen sein. Auch gilt es, die STI-Messung weiterer echokardiographischer Messverfahren der STIs hinsichtlich ihrer Wiederholbarkeit zu überprüfen. Beispielsweise fehlen Studien unserer Art für STI-Untersuchungen im M-Mode, mit der Speckle-Tracking-Analyse oder mit Hilfe der Tissue-Doppler-imaging-Methode, die ebenfalls in den letzten Jahren zur STI-Ermittlung genutzt wurden.<sup>48</sup>

Nicht betrachtet wurden, wegen des Informationsverlustes gegenüber der Einzelbetrachtung von PEP und LVET, die RIs und der Kappa Koeffizient der WR (PEP/LVET). In zukünftigen Studien könnte dies ein Untersuchungsschwerpunkt sein.

Schwierigkeiten bestehen bezüglich der Definition eines erfahrenen Untersuchers. Dahingehend wären zentrumsinterne oder sogar zentrumsübergreifende regelmäßige standardisierte Weiterbildungen wünschenswert. Auch könnte eine weitere Studie mit



Reifgeborenen erfolgen. Möglicherweise wäre hierbei eine weitere Steigerung der Reproduzierbarkeit der Messungen aufgrund besserer Schallbedingungen, längerer STIs etc. möglich.

Langfristig wird in größeren Studien untersucht werden müssen, ob sich die STI Erhebung gegenüber neueren Methoden, wie beispielsweise der seit einigen Jahren populären Speckle-strain-Echokardiographie, durchsetzen kann.

Die vorliegende Studie hat zu wichtigen Erkenntnissen bezüglich der klinischen Anwendbarkeit der STI-Erhebung geführt. Die wichtigste Erkenntnis für den klinischen Alltag ist, dass nur die dopplersonographische VET-Ermittlung eine ausreichende Reproduzierbarkeit und damit diagnostische Verlässlichkeit gewährleistet. Die PEP hat im optimalen Fall eine moderate Verlässlichkeit geboten. Zudem zeigten intraindividuell wiederholte Messungen in der Regel eine höhere Übereinstimmung. Auch ein 5-Minuten-APGAR  $\geq 8$ , sowie ein Geburtsgewicht  $\geq 1500\text{g}$  scheinen die Übereinstimmung der Messungen positiv zu beeinflussen. Umgekehrt verringern ein 5-Minuten-APGAR  $< 8$  und ein GG  $< 1500\text{g}$ , sowie Frühgeburtlichkeit die Genauigkeit. Es gilt zu erwähnen, dass eine Ko-Abhängigkeit zwischen den Faktoren besteht.

Entsprechend kann die VET-Messung als Standarddiagnostikum im Alltag auf neonatologischen Stationen zur Beurteilung des hämodynamischen Status Neugeborener in Betracht gezogen werden, insbesondere, wenn die Individuen ein GG  $\geq 1500\text{g}$  und einen 5-Minuten-APGAR-Wert  $\geq 8$  aufweisen. Derzeit gibt es dahingehend noch keine generellen Empfehlungen.

## 8. Abbildungsverzeichnis

Abbildung 1 Skizze PEP bzw. VET-Messung mit Hilfe von Flussprofil (A) und EKG (B); siehe vergleichend Abbildung 2 unter 4.3 .....	4
Abbildung 2 Messung STI über Aortenklappe .....	15
Abbildung 3 Messung STI über Pulmonalklappe .....	15
Abbildung 4 Messung STI über Aortenklappe bei bestehendem PDA .....	15
Abbildung 5 Beispielhafte Darstellung eines Bland et Altmann Diagrammes.....	17
 Diagramm 1 .....	 30
Diagramm 2 .....	31

## 9. Tabellenverzeichnis

Tabelle 1 Interpretation Kappa Werte nach Landis und Koch .....	21
Tabelle 2 Beispielhafte Darstellung einer generierten Kreuztabelle mit SPSS im Rahmen der Datenauswertung anhand des Kappa-Koeffizienten nach Cohen .....	22
Tabelle 3 zugehörig zu Tab. 2, berechneter Kappa-Wert nach $K = (p_o - p_e) / (1 - p_e)$ .....	23
Tabelle 4 Übersicht Basisdaten.....	25
Tabelle 5 Übersicht Untersuchungsvoraussetzungen.....	26
Tabelle 6 Differenzen der Parameter zu den Untersuchungsvoraussetzungen.....	27
Tabelle 7 Mittelwerte der STIs aus allen 3 Messungen der Untersuchungen .....	28
Tabelle 8 Mittelwerte der STIs/WRs aller betrachteten Untergruppen der gesamten Kohorte .....	28
Tabelle 9 Repeatability coefficient/Repeatability index der Differenzen aller Messwerte und gefiltert nach Erfahrungslevel .....	33
Tabelle 10 Repeatability coefficient/Repeatability index der Differenzen aller Messwerte gefiltert nach SST/GG/APGAR .....	35
Tabelle 11 Repeatability coefficient/Repeatability index der Differenzen aller Messwerte gefiltert nach KA-Gabe/PDA.....	36
Tabelle 12 Repeatability coefficient/Repeatability index der Differenzen aller Messwerte gefiltert nach Geschlecht .....	37
Tabelle 13 Repeatability coefficient/Repeatability index der Differenzen aller Messwerte gefiltert nach O <sub>2</sub> -Gabe/Beatmung .....	37
Tabelle 14 Repeatability coefficient/Repeatability index der Differenzen der Messwerte und gefiltert nach Erfahrungslevel, HF und VAS gleichbleibend .....	40
Tabelle 15 Repeatability coefficient/Repeatability index der Differenzen der Messwerte gefiltert nach Geschlecht, HF und VAS gleichbleibend .....	41
Tabelle 16 Repeatability coefficient/Repeatability index der Differenzen der Messwerte gefiltert nach SST/GG/APGAR, HF und VAS gleichbleibend .....	43
Tabelle 17 Repeatability coefficient/Repeatability index der Differenzen der Messwerte gefiltert nach KA-Gabe/PDA, HF und VAS gleichbleibend .....	44
Tabelle 18 Repeatability coefficient/Repeatability index der Differenzen der Messwerte gefiltert nach O <sub>2</sub> -Gabe/Beatmung, HF und VAS gleichbleibend .....	45
Tabelle 19 Cohens Kappa Koeffizient gesamte Kohorte und gefiltert nach Expertise.....	47
Tabelle 20 Cohens Kappa Koeffizient gefiltert nach Geschlecht .....	48
Tabelle 21 Cohens Kappa Koeffizient gefiltert nach SST/GG/APGAR und LT .....	51
Tabelle 22 Cohens Kappa Koeffizient gefiltert nach O <sub>2</sub> -Gabe und Beatmung.....	52
Tabelle 23 Cohens Kappa Koeffizient gefiltert nach KA-Gabe und PDA .....	53
Tabelle 24 Cohens Kappa Koeffizient bei gleichbleibender HF/VAS und gefiltert nach Expertise.....	55
Tabelle 25 Cohens Kappa Koeffizient bei gleichbleibender HF/VAS gefiltert nach Geschlecht .....	56
Tabelle 26 Cohens Kappa Koeffizient bei gleichbleibender HF/VAS gefiltert nach SST/GG/APGAR und LT.....	58
Tabelle 27 Cohens Kappa Koeffizient bei gleichbleibender HF/VAS gefiltert nach O <sub>2</sub> -Gabe und Beatmung .....	59
Tabelle 28 Cohens Kappa Koeffizient bei gleichbleibender HF/VAS gefiltert nach KA-Gabe und PDA.....	60

Tabelle 29 Betrachtung der Ausreißer der generierten Bland et Altman Diagramme für die linksventrikulären STIs.....	83
Tabelle 30 Betrachtung der Ausreißer der generierten Bland et Altman Diagramme für die rechtsventrikulären STIs .....	84

## 10. Literaturverzeichnis

- 1 Oh JK, Tajik J. The return of cardiac time intervals. *Journal of the American College of Cardiology*. 2003;42(8):1471-1474. doi:10.1016/S0735-1097(03)01036-2.
- 2 Weissler AM, Harris W, Schoenfeld C. Systolic Time Intervals in Heart Failure in Man. *Circulation*. 1968;37(2):149-159. doi:10.1161/01.CIR.37.2.149.
- 3 Barrington KJ. Common hemodynamic problems in the neonate. *Neonatology*. 2013;103(4):335-340. doi:10.1159/000349933.
- 4 Lopaschuk GD, Collins-Nakai RL, Itoi T. Developmental changes in energy substrate use by the heart. *Cardiovasc Res*. 1992;26(12):1172-1180. doi:10.1093/cvr/26.12.1172.
- 5 Teitel DF, Sidi D, Chin T, Brett C, Heymann MA, Rudolph AM. Developmental changes in myocardial contractile reserve in the lamb. *Pediatr Res*. 1985;19(9):948-955. doi:10.1203/00006450-198509000-00017.
- 6 Cayabyab R, McLean CW, Seri I. Definition of hypotension and assessment of hemodynamics in the preterm neonate. *J Perinatol*. 2009;29 Suppl 2:S58-62. doi:10.1038/jp.2009.29.
- 7 Noori S, Friedlich P, Seri I, Wong P. Changes in Myocardial Function and Hemodynamics after Ligation of the Ductus Arteriosus in Preterm Infants. *J Pediatr*. 2007;150(6):597-602. doi:10.1016/j.jpeds.2007.01.035.
- 8 Perlman JM, Cilio MR, Polin RA, eds. *Neurology: Neonatology questions and controversies*. Third edition. Philadelphia, PA: Elsevier; 2019.
- 9 Munro MJ, Walker AM, Barfield CP. Hypotensive extremely low birth weight infants have reduced cerebral blood flow. *Pediatrics*. 2004;114(6):1591-1596. doi:10.1542/peds.2004-1073.
- 10 Nagdyman N, Ewert P, Peters B, Miera O, Fleck T, Berger F. Comparison of different near-infrared spectroscopic cerebral oxygenation indices with central venous and jugular venous oxygenation saturation in children. *Paediatr Anaesth*. 2008;18(2):160-166. doi:10.1111/j.1460-9592.2007.02365.x.
- 11 Matcher SJ, Cooper CE. Absolute quantification of deoxyhaemoglobin concentration in tissue near infrared spectroscopy. *Phys Med Biol*. 1994;39(8):1295-1312. doi:10.1088/0031-9155/39/8/008.
- 12 Weisz DE, Jain A, Ting J, McNamara PJ, El-Khuffash A. Non-invasive cardiac output monitoring in preterm infants undergoing patent ductus arteriosus ligation: a comparison with echocardiography. *Neonatology*. 2014;106(4):330-336. doi:10.1159/000365278.
- 13 Jana Hartwich. *Vergleich von Impedanzkardiographie und Echokardiographie bei der linksventrikulären Schlagvolumenmessung bei Neugeborenen*. [Monographie-Dissertation]. Hamburg: medizinische Fakultät der Universität Hamburg; 2016.
- 14 Gillebert TC, van de Veire N, Buyzere ML de, Sutter J de. Time intervals and global cardiac function. Use and limitations. *Eur Heart J*. 2004;25(24):2185-2186. doi:10.1016/j.ehj.2004.10.017.
- 15 Robel-Tillig E. *Dopplersonographie in der Neonatologie*. 2nd ed. Berlin, Heidelberg: Springer Berlin Heidelberg; 2017.
- 16 Reant P, Dijos M, Donal E, et al. Systolic time intervals as simple echocardiographic parameters of left ventricular systolic performance: correlation with ejection fraction and longitudinal two-dimensional strain. *Eur J Echocardiogr*. 2010;11(10):834-844. doi:10.1093/ejehocardi/jeq084.

- 17 Alverson DC, Eldridge MW, Johnson JD, Aldrich M, Angelus P, Berman W. Noninvasive measurement of cardiac output in healthy preterm and term newborn infants. *Am J Perinatol*. 1984;1(2):148-151. doi:10.1055/s-2007-999991.
- 18 van de Werf F, Piessens J, Kesteloot H, Geest H de. A comparison of systolic time intervals derived from the central aortic pressure and from the external carotid pulse tracing. *Circulation*. 1975;51(2):310-316.
- 19 Robel-Tillig E, Knüpfer M, Pulzer F, Vogtmann C. Dopplersonographische Befunde in der Diagnostik des hämodynamisch wirksamen Ductus arteriosus. *Z Geburtshilfe Neonatol*. 2002;206(2):51-56. doi:10.1055/s-2002-30137.
- 20 Scalzi M de, Leonardis V de, Citi S, Cinelli P. Relationship between systolic time intervals and arterial blood pressure. *Clin Cardiol*. 1986;9(11):545-549. doi:10.1002/clc.4960091104.
- 21 Tarazi R.C., Frolich E.D., Dustan H.P. Left atrial abnormality and ventricular pre-ejection period in hypertension. *Dis. Chest*. 1969;55(214).
- 22 Chirife R, Ruiz GA, Gayet E, et al. The systolic index: a noninvasive approach for the assessment of cardiac function: implications for patients with DDD and CRT devices. *Pacing Clin Electrophysiol*. 2013;36(10):1284-1293. doi:10.1111/pace.12200.
- 23 Lewis RP, Rittogers SE, Froester WF, Boudoulas H. A critical review of the systolic time intervals. *Circulation*. 1977;56(2):146-158. doi:10.1161/01.cir.56.2.146.
- 24 Garrod AH. On some points connected with the circulation of the blood, arrived at from a study of the sphygmograph-trace. *Abstr. Pap. Printed Phil. Trans. R. Soc. Lond*. 1875;23(156-163):140-151. doi:10.1098/rspl.1874.0019.
- 25 Yared K, Noseworthy P, Weyman AE, McCabe E, Picard MH, Baggish AL. Pulmonary artery acceleration time provides an accurate estimate of systolic pulmonary arterial pressure during transthoracic echocardiography. *J Am Soc Echocardiogr*. 2011;24(6):687-692. doi:10.1016/j.echo.2011.03.008.
- 26 Pereira SS, Kempley ST, Shah DK, Morris JK, Sinha AK. Early echocardiography does not predict subsequent treatment of symptomatic patent ductus arteriosus in extremely preterm infants. *Acta Paediatr*. 2018;107(11):1909-1916. doi:10.1111/apa.14361.
- 27 Kulkarni M, Gokulakrishnan G, Price J, Fernandes CJ, Leeflang M, Pammi M. Diagnosing significant PDA using natriuretic peptides in preterm neonates: a systematic review. *Pediatrics*. 2015;135(2):e510-25. doi:10.1542/peds.2014-1995.
- 28 The Vermont-Oxford Trials Network: very low birth weight outcomes for 1990. Investigators of the Vermont-Oxford Trials Network Database Project. *Pediatrics*. 1993;91(3):540-545.
- 29 Hamrick SEG, Hansmann G. Patent ductus arteriosus of the preterm infant. *Pediatrics*. 2010;125(5):1020-1030. doi:10.1542/peds.2009-3506.
- 30 Richards J, Johnson A, Fox G, Campbell M. A second course of ibuprofen is effective in the closure of a clinically significant PDA in ELBW infants. *Pediatrics*. 2009;124(2):e287-93. doi:10.1542/peds.2008-2232.
- 31 Schwarz CE, Preusche A, Baden W, Poets CF, Franz AR. Repeatability of echocardiographic parameters to evaluate the hemodynamic relevance of patent ductus arteriosus in preterm infants: a prospective observational study. *BMC Pediatr*. 2016;16. doi:10.1186/s12887-016-0552-7.
- 32 Prescott S, Keim-Malpass J. Patent Ductus Arteriosus in the Preterm Infant: Diagnostic and Treatment Options. *Adv Neonatal Care*. 2017;17(1):10-18. doi:10.1097/ANC.0000000000000340.

- 33 Alverson DC, Eldridge MW, Johnson JD, et al. Effect of patent ductus arteriosus on left ventricular output in premature infants. *J Pediatr*. 1983;102(5):754-757.
- 34 Behrman RE, Martin CG, Snider AR, Katz SM, Peabody JL, Brady JP. Abnormal cerebral blood flow patterns in preterm infants with a large patent ductus arteriosus. *J Pediatr*. 1982;101(4):587-593. doi:10.1016/S0022-3476(82)80715-4.
- 35 Hermes-DeSantis ER, Clyman RI. Patent ductus arteriosus: pathophysiology and management. *J Perinatol*. 2006;26 Suppl 1:S14-8; discussion S22-3. doi:10.1038/sj.jp.7211465.
- 36 Koch J, Hensley G, Roy L, Brown S, Ramaciotti C, Rosenfeld CR. Prevalence of spontaneous closure of the ductus arteriosus in neonates at a birth weight of 1000 grams or less. *Pediatrics*. 2006;117(4):1113-1121. doi:10.1542/peds.2005-1528.
- 37 Dani C, Mosca F, Cresi F, et al. Patent ductus arteriosus in preterm infants born at 23-24 weeks' gestation: Should we pay more attention? *Early Hum Dev*. 2019;135:16-22. doi:10.1016/j.earlhumdev.2019.06.002.
- 38 Zonnenberg I, Waal K de. The definition of a haemodynamic significant duct in randomized controlled trials: a systematic literature review. *Acta Paediatr*. 2012;101(3):247-251. doi:10.1111/j.1651-2227.2011.02468.x.
- 39 Frederik Funke. *Vergleich visueller Analogskalen mit Kategorienskalen in Offline- und Onlinedesign*. [Magisterarbeit der]. Gießen: Justus-Liebig-Universität; 2004.
- 40 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*. 1986;1(8476):307-310.
- 41 McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276-282.
- 42 Halliday H, Hirschfeld S, Riggs T, Liebman J, Fanaroff A. Echographic ventricular systolic time intervals in normal term and preterm neonates. *Pediatrics*. 1978;62(3):317-321.
- 43 Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics*. 1977;33(1):159. doi:10.2307/2529310.
- 44 Cantor A, Wanderman KL, Karolevitch T, Ovsyshcher I, Gueron M. Systolic time intervals in children: normal standards for clinical use. *Circulation*. 1978;58(6):1123-1129. doi:10.1161/01.cir.58.6.1123.
- 45 Lang-Jensen T. Blood flow velocity and systolic time intervals measured by pulsed Doppler ultrasound: reproducibility of measurements. *Cardiovasc Res*. 1987;21(8):582-586. doi:10.1093/cvr/21.8.582.
- 46 Wanderman KL, Hayek Z, Ovsyshcher I, et al. Systolic time intervals in adolescents. Normal standards for clinical use and comparison with children and adults. *Circulation*. 1981;63(1):204-209. doi:10.1161/01.cir.63.1.204.
- 47 Hirschfeld S, Meyer R, Schwartz DC, Korfhagen J, Kaplan S. Measurement of right and left ventricular systolic time intervals by echocardiography. *Circulation*. 1975;51(2):304-309. doi:10.1161/01.cir.51.2.304.
- 48 Tavakolian K. Systolic Time Intervals and New Measurement Methods. *Cardiovasc Eng Technol*. 2016;7(2):118-125. doi:10.1007/s13239-016-0262-1.
- 49 Hedvall G. Systolic time intervals in children with heart disease. *Acta Paediatr Scand*. 1983;72(2):185-189. doi:10.1111/j.1651-2227.1983.tb09694.x.
- 50 van Dijk APJ, Hopman JCW, Klaessens J, van der Werf T, Daniëls O. The intra- and inter-operator variability of Echo-Doppler measurements for noninvasive pulmonary arterial pressure determination in congenital heart disease. *European Journal of Ultrasound*. 1996;4(2):91-98. doi:10.1016/0929-8266(95)00179-4.

- 51** Dehkordi P, Khosrow-Khavar F, Di Rienzo M, et al. Comparison of Different Methods for Estimating Cardiac Timings: A Comprehensive Multimodal Echocardiography Investigation. *Front Physiol.* 2019;10:1057. doi:10.3389/fphys.2019.01057.
- 52** Carvalho P, Paiva RP, Couceiro R, et al. Comparison of systolic time interval measurement modalities for portable devices. *Conf Proc IEEE Eng Med Biol Soc.* 2010;2010:606-609. doi:10.1109/IEMBS.2010.5626642.



## 11. Anhang

Tabelle 29 Betrachtung der Ausreißer der generierten Bland et Altmann Diagramme für die linksventrikulären STIs

	Ausreißer	Geschlecht	KA	PDA	SST (d)	O <sub>2</sub>	GG (g)
<b>LPEP1-2</b>	> +2SD	♂	nein	nein	<259	ja	< 1500
		♂	nein	nein	<259	ja	< 1500
	< -2SD	♂	ja	ja	≥259	nein	≥ 1500
<b>LPEP1-3</b>	> +2SD	♂	ja	ja	<259	ja	< 1500
		♀	nein	ja	≥259	nein	≥ 1500
	< -2SD	♀	nein	nein	<259	ja	< 1500
<b>LVET1-2</b>	> +2SD	♂	ja	ja	<259	ja	< 1500
	< -2SD	♂	nein	ja	<259	ja	< 1500
<b>LVET1-3</b>	> +2SD	♂	ja	ja	<259	ja	< 1500
		♂	nein	nein	<259	nein	≥ 1500
	< -2SD	♂	ja	nein	<259	ja	≥ 1500
		♂	nein	ja	<259	nein	≥ 1500

LPEP1-2 = Messung 1 durch Untersucher 1 - Messung 2 durch Untersucher 1 der LPEP; LPEP1-3 = Messung 1 durch Untersucher 1 - Messung 3 durch Untersucher 2 der LPEP; entsprechend für LVET; KA = Katecholamine, PDA = persist. Ductus arteriosus botalli, SST = Schwangerschaftstage; O<sub>2</sub> = Sauerstoff, GG = Geburtsgewicht

Tabelle 30 Betrachtung der Ausreißer der generierten Bland et Altmann Diagramme für die rechtsventrikulären STIs

	Ausreißer	Geschlecht	KA	PDA	SST (d)	O <sub>2</sub>	GG (g)
<b>RPEP1-2</b>	< -2SD	♂	nein	ja	<259	nein	≥ 1500
		♀	ja	nein	<259	ja	≥ 1500
<b>RPEP1-3</b>	> +2SD	♂	nein	ja	<259	ja	< 1500
		♀	nein	nein	<259	nein	< 1500
		♀	ja	ja	<259	ja	≥ 1500
	< -2 SD	♀	nein	nein	≥259	nein	≥ 1500
		♂	nein	nein	<259	nein	< 1500
		♂	nein	nein	<259	ja	< 1500
		♂	nein	ja	<259	nein	< 1500
		♀	nein	nein	<259	nein	< 1500
		♀	ja	nein	<259	ja	≥ 1500
<b>RVET1-2</b>	> +2SD	♂	nein	nein	<259	ja	< 1500
		♂	ja	ja	<259	nein	< 1500
		♂	ja	ja	<259	ja	< 1500
	< -2SD	♂	ja	nein	<259	ja	< 1500
		♀	nein	nein	<259	ja	< 1500
<b>RVET1-3</b>	> +2SD	♂	ja	ja	<259	ja	< 1500
		♂	ja	ja	<259	nein	< 1500
		♀	ja	nein	<259	nein	< 1500
	< -2SD	♂	nein	nein	<259	ja	< 1500
		♂	nein	nein	<259	ja	< 1500

RPEP1-2 = Messung 1 durch Untersucher 1 - Messung 2 durch Untersucher 1 der RPEP; RPEP1-3 = Messung 1 durch Untersucher 1 - Messung 3 durch Untersucher 2 der RPEP; entsprechend für RVET; KA = Katecholamine, PDA = persist. Ductus arteriosus botalli, SST = Schwangerschaftstage; O<sub>2</sub> = Sauerstoff, GG = Geburtsgewicht

## 12. Selbstständigkeitserklärung

### Erklärung über die eigenständige Abfassung der Arbeit

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne unzulässige Hilfe oder Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Ich versichere, dass Dritte von mir weder unmittelbar noch mittelbar eine Vergütung oder geldwerte Leistungen für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen, und dass die vorgelegte Arbeit weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt wurde. Alles aus anderen Quellen und von anderen Personen übernommene Material, das in der Arbeit verwendet wurde oder auf das direkt Bezug genommen wird, wurde als solches kenntlich gemacht. Insbesondere wurden alle Personen genannt, die direkt an der Entstehung der vorliegenden Arbeit beteiligt waren. Die aktuellen gesetzlichen Vorgaben in Bezug auf die Zulassung der klinischen Studien, die Bestimmungen des Tierschutzgesetzes, die Bestimmungen des Gentechnikgesetzes und die allgemeinen Datenschutzbestimmungen wurden eingehalten. Ich versichere, dass ich die Regelungen der Satzung der Universität Leipzig zur Sicherung guter wissenschaftlicher Praxis kenne und eingehalten habe.

07.09.2020

Datum



.....

Unterschrift

## Danksagung

Auf der letzten Seite dieser Arbeit verdienen die Menschen Dank, die auf ganz vielfältige Weise zur Fertigstellung dieser Arbeit beigetragen haben.

An erster Stelle möchte ich mich bei meinem Betreuer Professor Dr. med. Matthias Knüpfer bedanken, der mir das entsprechende Forschungsthema anvertraut hat, sodass diese Dissertationsarbeit entstehen konnte. Für Ihren fortwährenden fachlichen und kompetenten Rat bei wissenschaftlichen Fragen, sowie für Ihre zu jeder Zeit ermutigende Unterstützung, bedanke ich mich in aller Form. Ihr ständiges und freundliches Interesse bezüglich meiner beruflichen Laufbahn hat mich stets motiviert.

Ein weiterer Dank gilt dem ärztlichen und pflegerischen Personal der Neonatologie der Universitätsklinik Leipzig, ohne das die Erhebung des umfangreichen Datensatzes und eine optimale Betreuung der Neugeborenen während der Untersuchungen nicht möglich gewesen wäre. Auch bei den Eltern der Kinder, die uns ihr Vertrauen und die Zustimmung zu unserer Forschungsarbeit schenkten, möchte ich mich herzlich bedanken.

Ein außerordentlicher Dank gilt auch Dr. Bettina Altmann. Ihr gewinnbringendes Mitwirken und Ihre kompetente Beratung hinsichtlich der statistischen Methodik der Studie bereicherten die Arbeit ungemein. Die entsprechende Einweisung in die Analytik der Reproduzierbarkeit klinischer Methoden durch Sie war ausgesprochen hilfreich. Vielen Dank für die allzeit freundliche Zusammenarbeit.

Des Weiteren möchte ich mich bei meinen Freunden bedanken, die mich stets emotional unterstützt haben und ihre eigenen Erfahrungen und hilfreiche Hinweise mit mir geteilt haben. An dieser Stelle seien besonders Anna, Sarah, Robert, Lars, Bianka, Benni, Michael und Johannes genannt. Ich danke euch für die unabhängige Betrachtung der Thematik, woraus unglaublich wertvolle Erkenntnisse entstanden.

Zuletzt möchte ich von ganzem Herzen meiner Familie danken, ohne deren Unterstützung die Anfertigung dieser Arbeit in vielerlei Hinsicht nicht möglich gewesen wäre. Es ist schön, diese besondere Etappe meines Lebens mit euch zu teilen.